Expression of the VP1 antigen from foot-and-mouth disease virus in a bacterial and plant-based expression system

By

Priyen Pillay

Submitted in partial fulfilment of the requirements for the degree

MAGISTER SCIENTIAE

In the Faculty of Natural and Agricultural Sciences

UNIVERSITY OF PRETORIA

Pretoria

Supervisor: **Prof KJ Kunert**

Co-supervisor: **Dr R Chikwamba**

April 2012

DECLARATION

I, Priyen Pillay, declare that the dissertation, which I hereby submit for the degree Magister Scientiae at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

Signature: _____

Date: _____

TABLE OF CONTENTS

ABSTRACT

The suitability of a plant-based transient expression system using the agro-infiltration technique was compared to an *Escherichia coli* (*E. coli*)-based expression system to produce the VP1 protein from Serotype O, South Korean strain, of the foot-and mouth disease virus (FMDV). The full-length VP1 coding sequence was expressed in *Escherichia coli* as a fusion protein and purified as a His-tagged VP1 fusion protein with a yield of 14 mg L^{-1} bacterial culture. For transient expression in tobacco, the VP1 coding sequence was cloned into binary vector pMYV497, containing a CTB (cholera toxin B subunit) signal peptide and SEKDEL ER retention signal, and transiently agro-infiltrated into non-transgenic *N. benthamiana* and transgenic *N. tabacum* plants constitutively expressing the rice cysteine protease inhibitor OC-I. A protein resembling VP1 was detected using immuno-blotting analysis in both *N. benthamiana* and OC-I *N. tabacum* plants seven days post agro-infiltration. Although a possible stabilizing effect on VP1 was found due to OC-I expression, protein yields were not significantly different between transformed OC-I and non-OC-I control plants. Also, simultaneous co-infiltration with a plasmid allowing additional transient OC-I expression did not significantly improve VP1 production. The average VP1 amount achieved in OC-I expressing plants was 0.75% of total soluble protein. Overall, this study has shown that transient VP1 expression in tobacco is possible, but requiring further optimization, and that OC-I might have a stabilizing effect against proteolytic degradation of VP1 during advanced stages of senescence in agro-infiltrated plants coinciding with peaks in protein expression.

DISSERTATION COMPOSITION

Chapter 1 of the dissertation describes the VP1 genetic and protein architecture. This chapter discusses the expression of VP1 and VP1-related proteins in bacterial, plant and viral expression systems and also highlights the function and influence of cysteine proteinase inhibitors in recombinant protein production and overall plant functionality. It also covers protein targeting and post-translational modifications. **Chapter 2** describes the materials and methods used in this study. This encompassed DNA molecular cloning of the *VP1* gene sequence into a bacterial expression vector. The VP1 gene was also cloned into a binary vector for plant expression using transient agro-infiltration. Protein purification techniques, SDS-PAGE and immune-blotting analysis were used to evaluate VP1 protein expression. Fluorometric cysteine protease and GUS activity assays were carried out with control and transformed *N. tabacum* plants. **Chapter 3** describes the results obtained for expression of the recombinant VP1 protein in *E. coli*, *N. benthamiana* and *N. tabacum*. **Chapter 4** finally discusses the results obtained and proposes future perspectives for recombinant protein production using the co-expression of cysteine protease inhibitors. The **Annexure** contains plasmid maps and sequence alignments generated in this study. The literature cited in this study is listed under the chapter entitled **References**.

ACKNOWLEDGEMENTS

I wish to thank the funding body NRF (National Research Foundation) for financially supporting me throughout my research. I wish to extend special thanks to FABI (Forestry, Agricultural and Biotechnology Institute and the University of Pretoria for providing me with a conducive environment in which to do research. I wish to acknowledge the supportive role of my family, friends and my supervisors. I wish to thank Professor Karl Kunert, Dr Urte Schlüter, Professor Moon Sik Yang, Dr Tae Geum Kim and Dr Rachel Chikwamba for their support and guidance throughout my research. I wish to thank my lab members Stefan van Wyk, Rosita Endah, Endale Gebre, Berhanu Fenta Amsula, Tsholofelo Kibido, Abigail Mashamba, Magdeleen du Plessis, Celia Martins, Rynhard Smit and Dr Juan Vorster from the Plant Molecular Physiology Lab, Plant Sciences Department for their friendship and guidance. I wish to thank my lab members Kim Mi Young, Kim Bang Geul, Choi Jae Gwon, Yu Hwa Young, Jang Seon Hui, Oh Sun Mi, Nguyen Duc Chung, Nguyen Xuan Huy from the Plant Molecular Biology Lab, Division of Biological sciences, Chonbuk National University, Jeonju, Republic of Korea for their guidance and supervision. I also wish to thank my friends from FABI for their friendship and support. I extend special thanks to my close friends Theogaren Moodley, Waheed Mahomed and close cousin Vashlyn Pillay for their support and friendship throughout my journey. I also wish to thank my close uncle Gopaul Pillay who always supported and advised me. I extend a special thanks to my mother, Thangamal Pillay, for her care and love and for providing me with a conducive environment for learning, my father Kandasamy Pillay, whose excellent advice and love has always been a great help for me, and my two wonderful sisters Priyadeshni and Deshree Pillay for their love, support and encouragement. I wish to also extend a special thanks to my extended family for their support and encouragement.

ABBREVIATIONS AND SYMBOLS

DNA	deoxyribonucleic acid
FMDV	Foot-and-mouth disease virus
FMD	Foot-and-mouth disease
E. coli	*Escherichia coli*
CTB	Cholera toxin B subunit
HR	Hypersensitive response
%	Percent
ER	Endoplasmic reticulum
SP	signal peptide
PTMs	post-translational modifications
ERAD	Endoplasmic reticulum associated degradation
CaMV	Cauliflower mosaic virus
CPMV	Cowpea mosaic virus
PPVs	protease precursor vesicles
TMV	Tobacco mosaic virus
OC-I	Oryzacystatin-I
PDI	protein disulfide isomerase
°C	Degree Celsius
w/v	weight per volume percent
LB	Luria broth
µg	Microgram
ng	Nanogram
mL	Millilitre(s)

g	Gram
L	Litre
M	Molar
mM	Millimolar
OD_{600}	Optical density at a wavelength of 600 nanometers
µM	Micromolar
O/N	overnight
x g	Relative centrifugal force
min	Minute(s)
µL	Microlitre
ng	Nanogram
PCR	Polymerase Chain Reaction
ddH_2O	Double Distilled Water
sec	Second
IPTG	Isopropyl β-D-1-thiogalactopyranoside
X-gal	Bromo-chloro-indoly-galactopyranoside
hr	Hour(s)
rpm	revolutions per minute
Ni-NTA	Nickel-nitrilotriacetic acid
TSP	Total soluble protein
nm	Nanometer
mg	Milligram
MUG	4-methyl umbelliferyl glucuronide
MU	4-methyl umbelliferone
Z-Phe-Arg-MCA	Z-Phe-Arg-7-amido-4-methylcoumarin hydrochloride

MCA	α-amino 4-methylcoumarine
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
mA	milliampere
V	Volts
NBT	Nitrotetrazolium blue chloride
BCIP	5-Bromo-4-chloro-3-indoly phosphate disodium salt
PMSF	phenylmethylsulfonyl fluoride
GUS	β-glucuronidase
dpi	day(s) post-infiltration
pi	post-infiltration
(+) OC-I	*N. tabacum* plants stably expressing oryzacystatin-I
(-) OC-I	*N. tabacum* plants not expressing oryzacystatin-I
PTGS	post-transcriptional gene silencing
VLPs	virus like particles

INDEX OF FIGURES

INDEX OF TABLES

1. INTRODUCTION

1.1 The Foot-and-mouth disease virus (FMDV)

1.1.1 History

The significance of Foot-and-mouth disease (FMD) is highlighted in its discovery in the early 1500s by a monk named Hieronymus Fracastorius (Jones, 1977). He observed an epidemic occurring in cattle in Verona, Italy. The disease ravaged the cattle industry for many centuries after its discovery long before its pathogenic agent was discovered. In the late nineteenth century two former students of Robert Koch, Friedrich Loeffler and Paul Frosch were charged with the investigation of the cause of this disease (Mahy, 2005a). Their work showed that the pathological agent was small enough to pass through the pores of a special filter which was impermeable to the smallest known bacteria at the time (Loeffler and Frosch, 1898). Their discovery heralded the first description of a viral disease in animals (Loeffler and Frosch, 1897). They had discovered the Foot-and-mouth disease virus (FMDV).

In the Korean context, the VP1 protein of serotype O of FMDV expressed in this study, has been responsible for the slaughter of many livestock in 1999 and 2000 (Grubman and Baxt, 2004). As early as the year 2000 and 2002, outbreaks of FMDV, Serotype O, PanAsia have still occurred in the Republic of South Korea making it a challenging virus to contend with (Oem *et al.*, 2004). Recent outbreaks during the 2010-2011 period have occurred in South Korea with the most serious case occurring in Andong city where 1.4 million pigs were buried alive in an effort to control the spread of the disease. For this reason, an effective vaccine is desperately needed to curb the number of deaths. Details regarding the host-range,

pathogenesis are reviewed by Alexandersen and Mowat (Alexandersen and Mowat, 2005), and details regarding the transmission and spread of the virus are previously reviewed (Mahy, 2005b).

1.1.2 Serotypes

To date, seven serological types exist within FMDV which are based on the nature of their cell surface antigens. They are: type O, type A, type C, SAT1, SAT2, SAT3 and Asia 1 (Mahy, 2005a). Using the UPGMA method, Samuel and Knowles (2001) were able identify eight genotypes within the type O serotype of FMDV using nucleotide sequences from the 3' end of the VP1 gene with a cut-off value of 15% nucleotide difference (Samuel and Knowles, 2001). They have also undertaken similar studies in other serotypes: A, C and Asia 1 (Knowles and Samuel, 2003). The VP1 gene from Serotype O, PanAsia strain of FMDV was the target gene being investigated further in this study (Oem *et al.*, 2004).

1.1.3 Virus genetic architecture

Acharya *et al.* (1989) elucidated the three-dimensional structure of the virus using X-ray crystallography (Acharya *et al.*, 1989). FMDV is part of the Picornaviridae family, and is the type species of the genus *Aphthovirus* (Grubman and Baxt, 2004). These viruses contain a single-stranded positive-sense genomic RNA encapsulated in a 30 nanometer icosahedral virion with no envelope (Mahy, 2005a). The virus particles are comprised of 60 copies of four unique viral-encoded capsid proteins (VP1, VP2, VP3 and VP4) and a single copy of the viral RNA genome (Belsham, 2005). A more in-depth analysis of the translation and replication machinery of FMDV has been previously reviewed (Belsham, 2005). The VP1

protein as well as the other three capsid proteins are the constituents of the natural empty capsid and once 60 copies of each protein are made, they assemble into virus particles (Abrams *et al.*, 1995).

1.1.4 Structure and function of the viral capsid protein VP1

The VP1 gene has 639 base pairs and is reverse transcribed from linear RNA (NCBI) into a cDNA. The VP1 DNA sequence is translated into 213 amino acids forming the VP1 structural protein which has a molecular weight of 24 kDa (Fry *et al.*, 2005). Some studies place the size of the VP1 protein at approximately 26 kDa (Lentz *et al.*, 2010), which is the predicted size based on sequence analysis of VP1 from FMDV O_1K (Kurz *et al.*, 1981). The three-dimensional structure of the VP1 protein is available in the Protein Data Bank (PDB ID 1FOD) and is represented in Figure 1.1. There are two potential glycosylation sites present in the VP1 amino acid sequence located in the center (Asn-Thr-Thr-Asn-Pro-Thr) of the VP1 amino acid sequence (Kurz *et al.*, 1981). VP1 itself is not a glycoprotein despite interaction with the cell surface receptors which are glycoproteins (Grubman and Baxt, 2004). In FMDV A_{12} strain 119ab, all major structural polypeptides are phosphorylated to certain degrees with VP1 phosphorylated to a lesser extent (Torre *et al.*, 1980). Phosphorylation seemingly provides signalling for proper cleavage of precursors or for conformational changes in the protein during virus assembly of the outer shell (Torre *et al.*, 1980). Based on *in silico* analysis, the theoretical pI of the VP1 protein is 9.32 (http://web.expasy.org/protparam/).

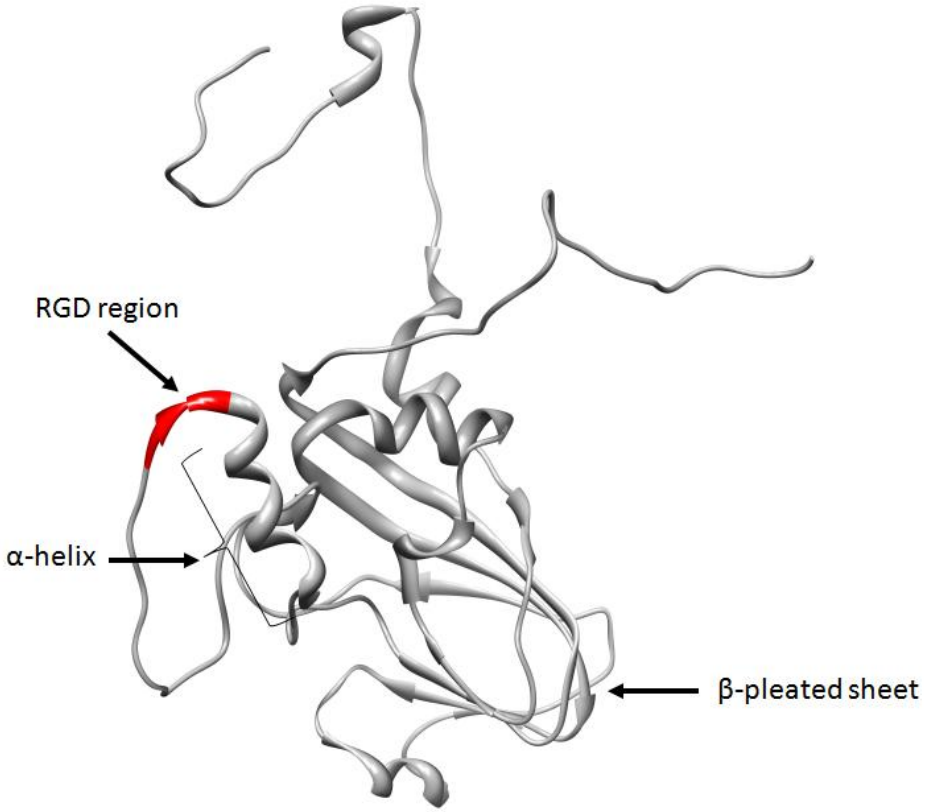


Figure 1.1 Three dimensional representation of the VP1 protein obtained from Protein Data Bank and modelled using USCF Chimera version 1.4 (PDB ID 1FOD). The immunogenically important RGD motif is indicated in red above. Spiral structures represent α-helices. Arrows represent β-pleated sheets.

Within FMDV serotype O_1K, removal of amino acid residues 138-154 causes a severe reduction in infectivity whereas infectivity is retained when a region between amino acid 200 and the C-terminus is cleaved which pins down cell attachment to a very small region represented by amino acids 138-154 (Strohmaier *et al.*, 1982). A structure within the VP1 protein called the GH-loop plays an important role in antibody and receptor recognition (Leippert *et al.*, 1997). Mutations within the Arg-Gly-Asp (RGD) sequence produced non-infectious viral particles indicating that the G-H loop of VP1 is a major FMDV antigenic site (Verdaguer *et al.*, 1995; Baranowski *et al.*, 2001). Further, the RGD (arginine-glycine-aspartic acid) loop attaches to host integrins on the cell surface and deletion of these three amino acids results in the loss of the cell attachment function (McKenna *et al.*, 1995). Removal of the RGD amino acid sequence (145 - 147 and/or amino acid residues 203 - 213 representing the C-terminal region of the VP1 protein) directly contributes to a loss in cell attachment (Fox *et al.*, 1989). There are similarities in the G-H loop structures of reduced type O virus and type C peptide which strongly points to the conserved nature of these structures across serotypes (Verdaguer *et al.*, 1995). Infection with FMDV induces cellular and humoral immune responses in infected hosts leading to the production of neutralizing antibodies. These are responsible for conferring protection from viral infection (Dus Santos and Wigdorovitz, 2005). The VP1 protein was selected for further investigation in our study due to its pivotal role as the major antigenic site for FMDV. Foot-and-mouth disease (FMD) is a highly contagious, acute vesicular disease affecting all cloven-hoofed animals, and is primarily responsible for large economic losses. Current measures to curb the spread of FMD include the control of animal movement, slaughter of infected and in-contact animals, disinfection and vaccination (Su *et al.*, 2007). Several newer expression systems to conventional vaccine development are emerging and many have significant advantages over conventional vaccines and are further discussed.

1.2 Production of heterologous proteins

1.2.1 Bacterial expression systems

With the rapid growth in the field of recombinant protein production, bacterial expression systems remain the most attractive due to low cost, high productivity, well-defined genetic characterization, wide variety of cloning vectors, mutant host strains as well as efficient use. Perhaps the most well-known bacterium used for heterologous protein production is *Escherichia coli* (*E. coli*) due to the sheer volume of knowledge available on its usage. An ubiquitous consideration for heterologous protein expression in any expression host is, in order to produce high levels of target protein, cloning the gene downstream of a well-characterized and regulated promoter is recommended. Some useful points to consider with regards to promoter are: strength, low basal expression levels (i.e. highly regulated), easily transferable to other *E. coli* strains, simple and cost-effective induction (Terpe, 2006). Generally, in bacteria, once recombinant proteins are over-expressed they accumulate either in the cytoplasm or periplasmic space between the outer and inner cell membrane of the bacterial cell. Thus a consideration for heterologous proteins containing disulfide bonds should be directing them to the periplasm instead of the highly reductive environment of the cytoplasm (Terpe, 2006). However, sometimes it is advantageous to direct proteins to the cytoplasm as it contains chaperon proteins which aid in the correct folding of the foreign protein (Choi and Lee, 2004). See Terpe (2006) for an extended overview of commonly used systems and the main features of bacterial promoters.

A major disadvantage that bacterial systems have is their inability to perform post-translational modifications, such glycosylation on proteins, in which case alternative

expression systems such as yeast, filamentous fungi, insect, mammalian cell cultures or plants should rather be employed (Terpe, 2006; Sahdev *et al.*, 2008). A trade-off with regards to high expression levels is the formation of inclusion bodies but there are many strategies available to overcome this challenge (Makrides, 1996). Optimization of expression levels requires a delicate balance of the combination of promoter strength and gene copy number to avoid the formation of inclusion bodies. The codons AGG, AGA, CUA, AUA, CCC, CGA are rarely used in *E. coli* and therefore may present problems for high-level expression (Terpe, 2006). This was further explored in our study as we were interested in expressing a plant-codon optimized gene sequence. Another major hindrance concerning heterologous protein expression in *E. coli* is the susceptibility to cell proteases resulting in low yields necessitating the exploration of other expression strategies or alternative host systems (Corchero *et al.*, 1996) such as the gram-positive Bacilli strains or plant-based expression in this study. An excellent example is the production of cholera toxin B subunit (CTB) in *Bacillus brevis* yielding 1400 mg per litre of culture (Ichikawa *et al.*, 1993). In contrast to *E. coli*, bacilli lack lipo-polysaccharides in the outer membrane and thus do not possess pyrogenic endotoxins that *E. coli* have and bacilli have a higher secretory potential readily secreting protein into the extracellular medium (Westers *et al.*, 2004). Another measure to inhibit protease action is the genetic conjugation of the target protein to a fusion counterpart which is also explored through this study. With regards to β-galactosidase fusion proteins, the placement site of the heterologous domain (either at the carboxy- or amino-end of β-galactosidase) was found to be highly influential in terms of solubility, proteolytic stability and final yield of full-length forms (Corchero *et al.*, 1996). The study with β-galactosidase fusions suggested that certain *E. coli* strains have lower degradation rates e.g. BL21 and thus are far more favourable to use. These observations are factored into the experimental design of our research with regards to which host is most suitable for heterologous protein

expression. Inclusion bodies do have some merit with regards to protection against proteolytic action as some proteins produced by bacterial systems sometimes form inclusion bodies which actually protect the recombinant protein from protease attack. Unfortunately, it may not necessarily produce the desired product as was the case with staphylococcal protein A fused with β-galactosidase (Hellebust *et al.*, 1989). See Table 1.1 for the general advantages and disadvantages of bacterial systems. Sahdev *et al.* (2007) stated "among the various expression systems employed for the over-production of proteins, bacteria still remains the favourite choice of a Protein Biochemist." It is encouraging to note that many counter strategies exist to combat the challenges that bacterial systems face (Sahdev *et al.*, 2008). The effectiveness of bacterial systems remains true and this applicability is explored further in this study.

1.2.2 Plant-based systems

Plants are fast becoming feasible and attractive expression systems for foreign proteins and using plants has now been explored for over 20 years (Rice *et al.*, 2005). Plants have several advantages over other expression systems. These include: the ability to correctly fold and assemble complex proteins and circumventing the need for utilizing the animal cells from diseased animals or animal-derived culture materials sources (Arntzen *et al.*, 2005; Doran, 2006). Other advantages are the available expertise in this field which is constantly growing, facilities for culture and storage of plant material (Arntzen *et al.*, 2005) and the ability to scale-up or scale-down depending on the production needs which are generally lower as opposed to other production systems (Twyman *et al.*, 2003; Streatfield, 2007). In light of these advantages, many plant bioreactors have been set up to increase the yields of protein

production (Doran, 2006), yield being the net-result of synthesis and breakdown (Stevens *et al.*, 2000).

1.2.2.1 *Transgenic plants*

Research into plant-based expression systems has mainly focused on obtaining balanced synthesis and proper assembly of the antigens or antibodies (Stevens *et al.*, 2000). The level of accumulation of the heterologous protein in transgenic plants is dependent upon many different factors. These factors also affect accumulation levels for transient expression using the agro-infiltration technique. These factors are: the transcription rate of the heterologous gene, mRNA stability, translatability of the foreign gene by the host's synthesizing machinery, codon usage (Perlak *et al.*, 1990) and finally the stability of the protein in the particular intra- or extracellular compartment to which it will be targeted (Wandelt *et al.*, 1992). These factors also determine yield, which has become the primary desire now that the field is developing and showing great potential. Researchers have shown the ability to produce various antigens and antibodies in plants but now the focus is shifting towards how much product can be produced and how production levels can be improved.

Unfortunately, the labour-intensive and time consuming process of generating large numbers of transgenic lines has been a hindrance for efficient functional analysis of heterologous proteins (Cazzonelli and Velten, 2006). In addition to this, transgenic plants suffer from being less stable under glasshouse conditions. The transgenic approach has previously yielded only low expression of FMDV-related antigens (Carrillo *et al.*, 1998; Wigdorovitz *et al.*, 1999a; Carrillo *et al.*, 2001) requiring extensive selection for high expressing transgenic lines (Dus Santos and Wigdorovitz, 2005). Foreign protein synthesis might also not keep pace with the

synthesis of the host leaf proteins effectively diluting the foreign protein out (Wandelt *et al.*, 1992). Wandelt *et al.* (1992) poignantly stated: "To achieve these levels of protein accumulation we anticipate that we will need to maximize transcription and translation as well as protein stability in transgenic plants." Despite these disadvantages, many studies (Tables 1.3 and 1.4) using transgenic plants have been pursued in an attempt to find an effective vaccine for the foot-and-mouth disease virus (FMDV).

1.2.2.2 *Transient expression (agro-infiltration)*

The advent of Agrobacterium-based plasmid vector system has allowed the transformation of a wide range of plant species based on a natural bacterial system to introduce foreign DNA sequences into the nuclear genome of plants. See Hellens *et al.* (2000) for a review of some of the systems that have been utilized for this purpose (Hellens *et al.*, 2000). The use of these vector systems were first intended for stable integration of the target gene into the plant genome (Zambryski *et al.*, 1983). Later on it was found that these systems can also be used in transient expression achieving very high expression levels (Kapila *et al.*, 1997). These evolved further through the advent of agro-infiltration which became the preferred method over stable transformation to assay gene silencing (Schöb *et al.*, 1997), conduct promoter analysis (Yang *et al.*, 2000) and investigate various genetic and physiological factors (Wroblewski *et al.*, 2005). Before generating stably transformed transgenic plants, transient gene expression analysis is particularly useful to rapidly confirm the correct assembly and retention of biological activity of the target protein (Johansen and Carrington, 2001; Rodríguez *et al.*, 2005). Transient expression is also greatly advantageous due to its efficiency and potential in that the number of genes that can be simultaneously expressed within the same cell could reach up to five or six, having great implications for the assembly

of complex multimeric proteins (Vaquero *et al.*, 1999). This particular characteristic is used within our study and the manner of its usage will be discussed further under the rationale section. One of the first transient expression studies was carried out on *Phaseolus vulgaris* leaves using vacuum infiltration. GUS activity post vacuum infiltration was comparable on three different days viz. day 2, 5 and 7 days post-infiltration and an interesting result was the variable GUS expression in different *Phaseolus* genotypes. In addition to this, the manner of incubation of infiltrated leaves also affected expression levels as shown by decreases in expression levels when incubated on solidified medium as opposed to wet filter paper. These observations were taken into account in the experimental design of this study. An advantage of the vacuum infiltration process was that no damage occurred to the plant cell wall (Kapila *et al.*, 1997). When tobacco plants were infiltrated, using vacuum infiltration, it was observed that the entire leaf surface area produced a blue colour. This was not the case in lettuce, tomato and Arabidopsis so there are limitations with vacuum infiltration (Wroblewski *et al.*, 2005). Nonetheless, using the Agrobacterium-mediated transformation method via vacuum infiltration, it became possible for bacteria to penetrate inner layers of plant tissue leading to T-DNA transfer to the mesophyll i.e. palisade and spongy parenchyma cells within the leaf tissue (Kapila *et al.*, 1997). In Petunia leaf discs, using transient Agrobacterium-mediated expression, leaves showed detectable GUS expression in histochemical assays 2 days post-inoculation with peaks in expression occurring between day 3 and 4 and then steadily decreasing over days 5-14. These peaks in expression were accounted for by the non-integration of T-DNA copies and are in accordance with the observations made by Wroblewski *et al.* (2005) using GUS staining and agro-infiltration. Quantitative GUS assays also reflected peaks in expression 3-4 days post-inoculation. These localized regions of GUS expression were also observed with several other species viz. *Nicotiana*, *Solanum* and *Malus* (Janssen and Gardner, 1990). In contrast to the decreases in GUS activity observed within

Petunia, increases in GUS activity were observed after the day 4 peak expression period in *Nicotiana sylvestris* plants persisting to day 10 (Schöb *et al.*, 1997).

Using agro-infiltration, *in vivo* expression analysis of promoters in agro-infiltrated tobacco leaves was possible after 2-3 days coinciding with peak expression (Wydro *et al.*, 2006). However, a possible disadvantage of the technique is that wounding of the leaf tissue is related to infiltration-associated wounding, agrobacterial infection activation of *PR* (pathogen-related) promoter genes (Yang *et al.*, 2000) and activation of pathogenesis-related proteins enhancing the host's defence response as shown by 20-fold increases in β-1,3-glucanase activity (Schöb *et al.*, 1997). Similarly, it was shown that the *vir* machinery of Agrobacterium is positively regulated at the transcriptional level when phenolic signal molecules (e.g. acetosyringone) are released from plant wounds (Rogowsky *et al.*, 1987). In addition, virulence and the plant's physiological condition itself influences the efficiency of transient assays. Agrobacterial virulence could be related to oncogene expression postulated to delay senescence in effect enhancing the efficacy of transient assays in some scenarios (Wroblewski *et al.*, 2005).

The utility of agro-infiltration has since been employed to express high-value proteins such as C5-1 murine antibody (D'Aoust *et al.*, 2009), antibodies for the human carcino-embryonic antigen (Vaquero *et al.*, 1999) and for this reason was employed to express the target VP1 protein within our study. Table 1.1 summarizes the advantages and disadvantages of stable and transient expression approached in plants as well as bacterial and viral expression systems. There are a number of plant-based recombinant drug products ranging from insulin to vaccines that are currently poised to go into the final stages of drug approval (Everett *et al.*, 2012).

Table 1.1 Summary of the advantages and disadvantages of various expression systems including transgenic and transient approaches.

Expression systems	**Bacterial**	**Viral**	**Plant**	
			Stable expression	**Transient expression (Agro-infiltration)**
Advantages	▪ Rapid growth rate ▪ Capacity for continuous fermentation ▪ Relatively low cost ▪ Protein enrichment option ▪ Protection against proteases ▪ Extremely high yields (Yin *et al.*, 2007)	▪ High transient expression levels ▪ Easy to manipulate viral genome ▪ System is simple and efficient ▪ Ability to confer immunity against some diseases (Canizares *et al.*, 2005)	▪ Release of GMOs are highly regulated preventing unwanted release of potentially harmful agents (Hellens *et al.*, 2000) ▪ Protein enrichment option (Delannoy *et al.*, 2008) ▪ Proteins of interest can be expressed at specific growth stages (Yin *et al.*, 2007) ▪ Heterologous proteins expressed can be localized to different organs ▪ Inexpensive compared to other systems (Yin *et al.*, 2007) ▪ Wide variety of vector systems available (Hellens *et al.*, 2000) ▪ Mediation of proper folding assembly in ER cellular compartment ▪ No risk of bloodborne pathogens and oncogenic sequences	▪ High expression levels ▪ Pervading of all leaf cell layers ▪ Gene expression measurement can be conducted shortly after delivery ▪ Unbiased expression due to the lack of positional effects ▪ Good for gene expression in recalcitrant plant species ▪ Highly efficient ▪ Minimal tissue culture manipulations ▪ High transformation efficiency ▪ Versatile, utility in a number of plant species (Kapila *et al.*, 1997) ▪ Reliable method (Yang *et al.*, 2000) ▪ Inexpensive (Vaquero *et al.*, 1999)

			▪ Suitable for upscaling production and protein purification (Vaquero *et al.*, 1999) ▪ Proven scalability, high production capacity, lower capital expenditures, decreased operating costs (Everett *et al.*, 2012)	
Disadvantages	▪ Protein product may be insoluble or misfolded ▪ Cannot perform the post-translational modifications ▪ Codon bias problems ▪ Degradation by proteases ▪ Instability of plasmids (Yin *et al.*, 2007) ▪ Contamination of endotoxins (Vaquero *et al.*, 1999)	▪ Small gene size can only be inserted into viral genome ▪ Foreign gene is not heritable ▪ Bio-containment concerns (Canizares *et al.*, 2005)	▪ Several-fold lower expression levels ▪ Biased expression due to positional effects ▪ Low transformation efficiency levels (Yin *et al.*, 2007) ▪ Only specific cells get transformed (Kapila *et al.*, 1997) ▪ Possibility of incomplete integration of target into plant genome (Hellens *et al.*, 2000) ▪ Labor-intensive, time-consuming (Vaquero *et al.*, 1999) ▪ Large amounts of variation in expression (Yang *et al.*, 2000) ▪ Different glycosylation patterns compared to animal-derived cell expression systems (James and Lee, 2001) ▪ Expensive (Vaquero *et al.*, 1999) ▪ Too tightly regulated preventing growth of field (Everett *et al.*, 2012)	▪ Activates pathogen-related defence pathways (Yang *et al.*, 2000) ▪ Elicits hypersensitive response (HR) which manifests in necrosis ▪ Only highly effective in a subset of plant species ▪ Provides a measurable phenotype only for a few days (Wroblewski *et al.*, 2005)

1.2.2.3 *Factors affecting expression of heterologous proteins in plants*

1.2.2.3.1 *Localization*

Protein production and accumulation can differ significantly between different plant cellular compartments and can be critical to the successful expression of potential vaccines candidates (Rice *et al.,* 2005). The directing of foreign proteins to specific cellular locations or organelles within the plant cell is one of the primary strategies for minimizing protein degradation and maximising yields (Doran, 2006; Benchabane *et al.,* 2008). Organelles have their own protease complement attuned to its enzymatic and physiochemical environment (Callis, 1995). These factors have also been taken into consideration in this study.

The structural characteristics of the recombinant protein of interest strongly dictate the choice of cellular compartment it should be directed to because post-translational modifications are imperative for the proper functioning of the foreign protein and its stability thereafter (Faye *et al.,* 2005). An example is the accumulation of human growth factor in either the cytosol or the apoplast which produced a toxic effect in *N. benthamiana* leaves. However, when the same protein was directed to the chloroplast, no negative effects were observed in the plant (Gils *et al.,* 2005). In light of an unpredictable reaction of the host plant to the introduction of recombinant proteins and the pleiotropic effects, it is advised that an empirical approach be taken by analysing all or some of the cellular compartments for protein accumulation (Benchabane *et al.,* 2008). Thus, many cellular compartments have been used for this purpose including the cytosol, chloroplasts and the endoplasmic reticulum (Daniell *et al.,* 2005; Henry, 2006; Li *et al.,* 2006; Petruccelli *et al.,* 2006). Figure 1.2 illustrates a variety of mechanisms for sub-cellular targeting via signal peptides.

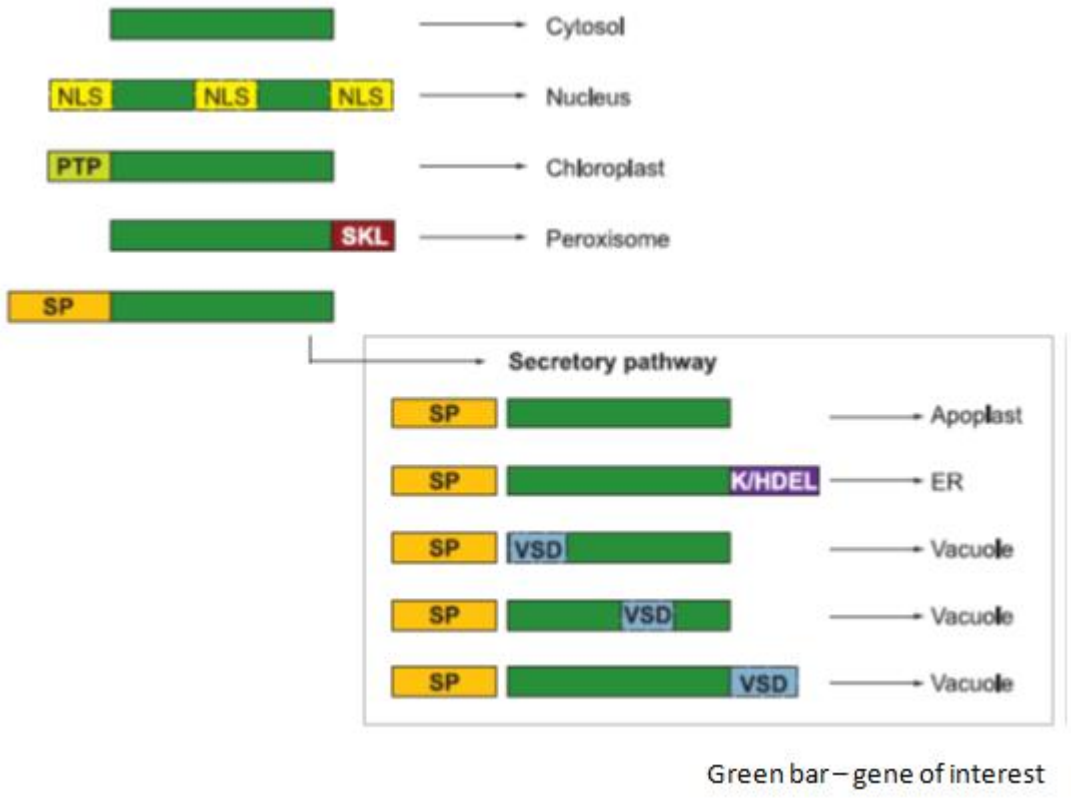


Figure 1.2 Schematic representation of sub-cellular targeting of recombinant proteins in plant cells. Recombinant proteins bearing an N-terminal signal peptide in their primary sequence will enter the cell secretory pathway via the endoplasmic reticulum (ER), and then travel through the Golgi system to be secreted in the apoplast (default pathway) or directed to the vacuole if a vacuolar sorting determinant (VSD) is present in the protein sequence. Proteins secreted into the ER can also be retained in this compartment by the grafting of an ER retention signal – the KDEL (or HDEL) tetra-peptide motif – at the C-terminus. Proteins with no signal peptide accumulate in the cytosol (default location) or migrate to specific organelles when an appropriate peptide signal is included in the transgene sequence. Peptide signals used recently in transgenic plant platforms include plastid (e.g. chloroplast) transit peptides (PTP), nuclear localization signals (NLS) and the tripeptide peroxisome target sequence serine-lysine-leucine (SKL) (adapted from Benchabane *et al*., 2008).

1.2.2.3.2 *Endoplasmic reticulum*

It has been shown in a number of studies that directing foreign proteins towards the endoplasmic reticulum, instead of the cytosol, produces a greater yield by minimizing proteolytic degradation (Conrad and Fiedler, 1998; Wydro *et al.*, 2006; Benchabane *et al.*, 2008). It is known that the endoplasmic reticulum has an increased accumulation of protein relative to the cytoplasm (Napier *et al.*, 1998; Ellgaard and Helenius, 2003). Suggested reasons for this are the presence of high amounts of chaperone proteins (proteins that assist with proper folding), a reduction in the level of plant proteases and the presence of disulfide isomerase which facilitates the formation of disulfide bonds in protein structures (Nuttall *et al.*, 2002; Ma *et al.*, 2005; Rice *et al.*, 2005). The ER does contain proteolytic proteases that could alter the structural integrity or heterogeneity of the protein which is a disadvantage for recombinant protein localisation. It has been shown that by localising foreign proteins to the ER, protein yield increases by 10-100 times (Hellwig *et al.*, 2004). However, the properties of the foreign protein of interest govern the appropriateness of targeting to the ER for protein accumulation due to post-translational processing that occurs more downstream of the ER along the secretory pathway in the Golgi apparatus (Gomord and Faye, 2004; Doran, 2006). Proteins can be directed to this organelle using signal peptides KDEL or HDEL (Ma *et al.*, 2003; Doran, 2006). The high yields, minimized proteolytic degradation and existence of chaperone proteins were determining factors within this study when deciding which cellular compartment to direct the VP1 protein towards.

1.2.2.3.3 *Cytosol*

If no signal peptide is attached to the foreign protein's sequence then it will not migrate out from the cytosol after its mRNA is translated (Benchabane *et al.*, 2008). In many cases recombinant proteins within the cytosol have low accumulation rates (Conrad and Fiedler, 1998) and for this reason was avoided within our study. The human growth hormone illustrates this point with accumulation levels of 0.01% of total soluble protein (Gils *et al.*, 2005). There are many reasons why the cytosol is not a suitable cellular compartment for recombinant protein targeting. This includes:- (i) modifications, such as glycosylation, which have an impact on the structural integrity of nascent and mature proteins present in the cytosol (Faye *et al.*, 2005); (ii) the negative potential within the cytosol which is unfavourable for the correct folding of proteins that require disulphide bonds and finally (iii) the effectiveness of the housekeeping activity of the ubiquitin-proteasome proteolytic pathway rendering protein accumulation very challenging as proteins that are improperly folded are recognized and degraded (Vierstra, 2003; Smalle and Vierstra, 2004).

1.2.2.3.4 *Apoplast*

The extracellular medium, also known as the apoplast is another locality where heterologous proteins can be expressed (Goulet *et al.*, 2011). However, it is undesirable due to its high proteolytic content (Callis, 1995) reinforcing the strategy of ER retention within our study. Nonetheless, there have been many recombinant proteins engineered for extracellular secretion over the last decade which includes the (i) LTB protein expressed in corn, (ii) human growth hormone expressed in *N. benthamiana* leaves, and the (iii) murine diagnostic

antibody C5-1 expressed in *N. benthamiana* leaves (Streatfield *et al.,* 2003; Hellwig *et al.,* 2004; Gils *et al.,* 2005; Goulet *et al.,* 2011).

1.2.2.3.5 *Vacuole*

Two types of vacuoles have been proposed to exist: (1) lytic vacuoles, which are rich in hydrolytic enzymes and are adapted to an acidic pH and (2) protein storage vacuoles, which are acidic to a lesser degree compared to lytic vacuoles but are adapted to protein storage (Robinson *et al.,* 2005). Lytic vacuoles are not suitable for heterologous protein retention due to their high protease content whereas protein storage vacuoles are more conducive for protein accumulation due to the more basic environment they present (Stoger *et al.,* 2005). These types of vacuoles are most abundant in seeds. The mechanism of targeting proteins to the vacuole are governed by amino acids within the actual protein primary sequence that act as sorting signals for directing proteins to the vacuole (Nakamura and Matsuoka, 1993; Neuhaus and Rogers, 1998). An example of vacuolar-targeting is the expression of a dog gastric lipase in transgenic tobacco plants (Gruber *et al.,* 2001) which is applicable owing to the nature and functioning environment of this particular protein. Vacuolar-targeting was avoided within this study due to the strong protease complement within the vacuole.

1.2.2.3.6 *Chloroplast*

A fairly new strategy of expressing foreign proteins in plastids of transgenic plants, such as the chloroplast, gives improved protein stability and protection against protease degradation (Daniell *et al.,* 2005). This system is advantageous for transgenic plants in that there are multiple copies of the chloroplast genome per plastid and there are many plastids per cell.

The consequences are an extraordinary increase in expression from 500- to 4000-fold as opposed to nuclear-based transgene expression (Daniell *et al.,* 2005). Using this system avoids challenges associated with positional effects and gene silencing that normally plague traditional nuclear-based transformations (Rice *et al.,* 2005). Transformation of the chloroplast has several advantages: including uniform expression rates, multiple copies of the transgene, expression of many genes from the same construct as well as low gene silencing. Another advantage is the maternal inheritance of chloroplast DNA in many plant species which minimizes transgene escape (Daniell *et al.,* 2002). The stroma of the chloroplast lends itself to post-translational modifications like disulphide bridging and multimerization (Daniell *et al.,* 2005) without the reliance on other more complex modifications like glycosylation that takes place in the cell secretory pathway. Studies in transgenic lettuce and tobacco lines show encouraging results where a cholera toxin B – pro-insulin fusion protein was expressed within the chloroplast (Ruhlman *et al.,* 2007). Tobacco chloroplasts are also a conductive environment for the high expression of the VP1 structural protein from the foot-and-mouth disease virus (Li *et al.,* 2006; Lentz *et al.,* 2010). As was mentioned earlier an improved production of the human growth hormone in *N. benthamiana* leaves was further observed by Gils *et al.* (2005). Despite the success in chloroplasts, the inability to perform more complex post-translational modifications, such as glycosylation and assembly processes including proper folding is disadvantageous for this organelle (Maliga, 2002; Rice *et al.,* 2005). Endogenous proteases are also present in the chloroplast which compromise the accumulation of proteins (Adam and Clarke, 2002). The activity of certain chloroplast proteases is illustrated in the rotavirus VP6 protein example where high accumulation rates of protein were observed in young tobacco leaves whereas in older leaves the amount of VP6 protein decreased which suggest their susceptibility to proteolytic degradation (Birch-Machin *et al.,* 2004). Chloroplast expression was avoided due to the disadvantages mentioned above.

1.2.2.4 *Signal peptides*

In both prokaryotic and eukaryotic cellular expression systems, proteins are only allowed to enter the secretory pathway if they possess a specific targeting signal peptide (SP) (von Heijne, 1990). In most cases the SP is only a transient extension to the amino terminus of the target protein and is subsequently removed by one of a small class of enzymes known as signal peptidases once its targeting function has been executed (von Heijne, 1990). SPs themselves are characterized by having three distinct domains which are (i) a positively charged amino-terminal region (n-region, 1-5 residues long) (ii) a centralized, hydrophobic region (h-region, 7-15 amino acid residues) and a polar carboxy-terminal domain (c-region, 3-7 amino acid residues) (von Heijne, 1990). The likely mechanism of functioning seems to be that the n- and h-regions are responsible for targeting the SP and protein of interest, with the c-region only needed for the removal of the SP from the mature polypeptide chain (von Heijne, 1990). It is postulated that the SP interacts primarily with receptor proteins and with the ER lipid bilayer itself (Schatz and Dobberstein, 1996). Of particular interest is the signal peptide from CTB (Cholera toxin B subunit) encoded by a 21-amino acid sequence, that is highly hydrophobic, upstream of the CTB amino acid sequence and is used to direct proteins to the secretory pathway (Gennaro and Greenaway, 1983). The native function of these signal peptides in *Vibrio cholerae* is facilitating the transport of each B subunit into the periplasmic space or to the outer membrane of the host cell (Gennaro and Greenaway, 1983). In plants, the 21-amino acid leader peptide of the CTB protein directs the newly synthesized CTB protein into the lumen of the ER (Arakawa *et al.,* 1997) which is the method of localization followed within this study. The origin of the signal peptide sequence is not critical as sequences from different organisms can still function (Ma and Hein, 1995) and information

encoded within each sequence motif dictates the specific functions of the signal sequence (Schatz and Dobberstein, 1996).

1.2.2.5 *Post-translational modifications*

One of the major advantages of plants over other production systems (yeast or bacterial expression platforms) is their ability to perform a host of the post-translational modifications (PTMs) which are essential for the bioactivity of recombinant therapeutic proteins (Gomord and Faye, 2004; Gomord *et al.*, 2010). The term PTM refers to a wide variety of modifications spanning covalent modifications yielding derivatives of individual amino-acid residues (e.g. glycosylation, phosphorylation, methylation, ADP-ribosylation, oxidation and glycation). This term also encompasses proteolytic processing of polypeptide backbones and non-enzymatic modifications such as deamidation, racemization and spontaneous changes with regards to protein conformation (Gomord and Faye, 2004).

With regards to N-linked glycosylation certain structural requirements must be met ensuring exposure of the Asn residue on the glycoprotein surface which may result in either fully glycosylated, partially glycosylated, or not glycosylated at all protein complexes (Vitale and Denecke, 1999). The process of N-linked glycosylation begins in the ER where there is a post-translational transfer of an oligosaccharide precursor $Glc_3Man_9GlcNAc_2$ specifically onto asparagine residues within the consensus sequence Asn-X-Ser/Thr (in which case X represents any amino acid except Pro, Ser represents serine and Thr represents threonine) (Gomord and Faye, 2004). It is postulated that the glycans attached to antibodies play a role in structural stability and protection against proteolytic degradation (Dwek, 1996). In producing glycosylated therapeutic proteins in plants, the production of non-immunogenic N-

glycans is desired. A strategy to prevent addition of immunogenic glycans to heterologous proteins involves the directing and storage of the therapeutic protein within the ER upstream of the Golgi cisternae. In this strategy, immunogenic glyco-epitopes are added to plant N-glycans (Gomord and Faye, 2004) or the amino acid site of glycan attachment is mutated (Rodríguez *et al.*, 2005). This has been illustrated by the expression of mouse/human chimeric IgG1 antibody fused with a C-terminal KDEL retention signal which exclusively presented high-mannose-type N-glycans with 6–9 mannose residues (Sriraman *et al.*, 2004). Mouse/human chimeric IgG1 antibody, which lacked the KDEL retention signal, underwent more complex modifications in the Golgi compartment producing N-glycans containing core-xylose and core-α(1,3)-fucose. This illustrated that the association of immunogenic N-glycans to heterologous proteins can be prevented by the fusion of ER retrieval signals to the protein of interest (Sriraman *et al.*, 2004). Another PTM to consider in plants is O-glycosylation. O-glycosylation occurs mainly on the hydroxyl groups of hydroxyproline, serine and threonine residues within hydroxyproline-rich glycoproteins (Gomord *et al.*, 2004).

Most therapeutic proteins, such as immuno-globulins, structural proteins and vaccines, are co-translationally inserted into the lumen of the ER, and subsequently transported via the Golgi to the lysosomal compartment, the extracellular matrix or the blood stream (Gomord and Faye, 2004) predisposing these proteins to complex PTMs. Thus PTM events and stability of heterologous proteins within the host plant is an important consideration as different hosts vary in their processing of heterologous proteins which is important from a functional standpoint (Menkhaus *et al.*, 2004) and the antigen's capacity to induce an effective immune response is dependent on its structure (Mikschofsky *et al.*, 2009). Folding and/or post-translational modification of heterologous proteins and proteins naturally residing

in the cytosolic compartment may differ if they are directed into the secretory pathway (Mikschofsky *et al.,* 2009). It is known for antibodies that after assembly, maturation takes place in the ER and Golgi apparatus in the absence of retention signals and subsequently these antibodies are excreted into the apoplastic space making them susceptible to a plethora of proteases (Stevens *et al.,* 2000). Therefore, the ultimate destination of the target protein within the plant after production is an important consideration (Menkhaus *et al.,* 2004). If, for example, the target protein contains disulphide bonds, it is then advantageous to direct it to the ER sub-cellular compartment (Mikschofsky *et al.,* 2009) as protein folding is aided by folding enzymes, such as protein disulfide isomerase (PDI), which catalyzes the formation of disulfide bridges (Vitale and Denecke, 1999). In order to effectively achieve this, the hexa-peptide SEKDEL was found to confer ER retention by marking proteins, e.g. lysozyme, that were normally secreted causing them to be retained (Munro and Pelham, 1987). The ER retention signal is a linear sequence as opposed to a complex three-dimensional structure (Munro and Pelham, 1987). Proteins that are retained in the ER accrete in structures, called protein bodies, which enhances their post-translational stability (Pueyo *et al.,* 1995). ER retention causes aggregation of many proteins that are normally soluble (Wandelt *et al.,* 1992). The value of retention within the ER sub-cellular compartment of antibodies has been reported by Sainsbury and Lomonossoff (2008) where an ER retention signal increased human anti-HIV 2G12 levels in *N. benthamiana* plants (Sainsbury and Lomonossoff, 2008). This was also shown with the fusion of the β-subunit of *Escherichia coli* heat-labile toxin to a SEKDEL ER retention signal, and stable transformation into *N. tabacum* L. cv TI560 plants where protein expression was enhanced 200-fold (Kang *et al.,* 2004) as well as with the retention of the CTB adjuvant to the ER compartment (Mikschofsky *et al.,* 2009). The study by Pan *et al.* (2008) has also illustrated the value of retaining proteins to the ER compartment where the structural poly-protein, P1-2A as well as 3C protease from FMDV serotype O,

strain China/1/99 was stably expressed in foliar tomato extract. Most importantly, ER retention has shown potential of increased stability and higher levels of accumulation in transgenic tobacco plants, expressing vicilin, accounted for by the lack of exposure of the heterologous protein to the proteolytic environment of the leaf vacuole (Wandelt *et al.*, 1992). In studies where the signal peptide (e.g. CTB signal peptide) was absent, no expression of the CTB adjuvant was observed, which was accounted for by degradation within the cytoplasm during or immediately after synthesis (Mikschofsky *et al.*, 2009).

There are disadvantages to ER retention that one should be consider. Polypeptides, which do not fold properly, are retained in the ER where many proteolytic pathways within the ER act on mis-folded products (Pagny *et al.*, 1999). Proteins that are mis-folded are in turn retro-translocated from the ER into the cytosol by the ER machinery destined for proteasomal degradation through ER associated degradation (ERAD) (Gomord *et al.*, 2010). The converse of retention is secretion of target proteins into the apoplastic space. However, as mentioned above, the apoplastic space contains a vast number of proteases to which the protein will be exposed to. Therefore, there is a trade-off when directing a protein through the entire secretory pathway (Stevens *et al.*, 2000). Due to the development of transient expression systems, the ER sub-cellular compartment has been harnessed to synthesize heterologous proteins for industrial and pharmaceutical purposes (Vitale and Denecke, 1999).

1.2.2.6 *Proteases*

There are six main classes of proteases: cysteine, glutamic, serine, threonine, aspartic and metallo – proteases (Fan and Wu, 2005). Proteolytic enzymes in plants can be divided into two categories: limited and unlimited proteolysis (Fan and Wu, 2005). Limited proteolysis

refers to proteases that cleave only one or a limited number of peptide bonds of a target protein leading to the activation or maturation of the formerly inactive protein whereas unlimited proteolysis refers to proteolytic action on proteins degrading them into their individual amino acid constituents (Fan and Wu, 2005). The major classes of peptidases (all enzymes that hydrolyze peptide bonds) can be further divided into two major groups that are endopeptidases and exopeptidases (Filho, 1992). The term "protease" encompasses both exopeptidases and endopeptidases while "proteinase" describes only endopeptidases (Barrett, 1986). Endopeptidases hydrolyze internal peptide bonds whereas exopeptidases hydrolyze either the N-terminal (amino-peptidases) or C-terminal (carboxy-peptidases) bonds exclusively (Filho, 1992; Fan and Wu, 2005; van der Hoorn, 2008).

The number of genes which encode for proteins involved in proteolytic processes is fairly large, estimated to be about 1900 in Arabidopsis (Schaller, 2004). In *N. tabacum* there are about 35 known or putative peptidases and 10 non-peptidase homologues (MEROPS database) which is considerably higher than *N. benthamiana* presenting a challenge for users of this species of tobacco for the expression of foreign proteins. In plants the overall yield of foreign proteins is governed by many factors, particularly the role played by plant proteases which bring about protein degradation (Doran, 2006; Benchabane *et al.*, 2008). The function of proteases covers some of the most important metabolic processes within the plant. This includes assembling, disassembling proteins and removing damaged, mis-folded or potentially harmful proteins (Vierstra, 1996; Fan and Wu, 2005) at various cues to ensure proper functioning (Filho, 1992). The function of selective proteolysis is the elimination of mis-folded proteins and recycling of amino acids from short-lived proteins by degrading them (Callis, 1995; Schaller, 2004) which has many implications for the expression of heterologous proteins in plants. Proteases may affect the structural integrity of recombinant

proteins at various stages, both *in planta* during protein expression and *ex planta* during extraction and downstream processing (Rivard *et al.*, 2006). Evidence of this is suggested by studies showing increases in mRNA levels of genes which encode for proteases involved in developmental and environmental changes that plants undergo during their lifetime (Callis, 1995; van der Hoorn, 2008). Protein turnover is also regulated by the ubiquitin/proteasome pathway which targets lysine residues (Buchanan-Wollaston, 1997; Vierstra, 2003). In plants, proteins that are to be degraded are first conjugated to many molecules of the polypeptide ubiquitin (Fan and Wu, 2005). There are many intracellular localities within the plant where proteolysis occurs and within these localities exists a plethora of proteolytic pathways (Callis, 1995). Proteases mainly exist both in the chloroplast and in the mitochondrial sub-cellular compartments (Adam *et al.*, 2001) as well as in the lysosome (Fan and Wu, 2005). According to Goulet *et al.* (2011), for the *Nicotiana* species, the majority of protease families are targeted to the secretory pathway were mostly aspartic proteases and cysteine proteases (papain-like) and to a lesser extent serine and metallo-proteases reside (Delannoy *et al.*, 2008; Goulet *et al.*, 2011). This trend was taken into account within this study.

One of main groups of proteases, the cysteine proteases, which were also investigated in more detail for their action in this study, are involved in many developmental processes. The active site for cysteine proteases are cysteine amino acid residues (Oliveira *et al.*, 2003) and during proteolysis, cysteine proteases use a catalytic cysteine as a nucleophile to achieve cleavage (van der Hoorn, 2008). Catalysis proceeds through the formation of a covalent intermediate involving a cysteine (Cys 25) and a histidine (His 159) residue (Fan and Wu, 2005).

Papain was the first cysteine protease to be discovered originating from the latex and fruit of *Carica papaya* (Drenth *et al.*, 1962; Light *et al.*, 1964). Papain-like enzymes are of great interest as they are involved protein degradation and nitrogen-mobilization during the processes of seed germination (Callis, 1995) as well as during leaf senescence (Ueda *et al.*, 2000). The Cathepsin L-like cysteine proteases, which are endo-peptidases (Oliveira *et al.*, 2003), and legumain are quite prevalent in whole *N. benthamiana* leaf protein extracts as compared to apoplast extracts suggesting an intracellular location for these enzymes (Goulet *et al.*, 2011). Previous work also shows that primarily aspartic, cysteine and serine peptidases occur in the apoplast of *N. tabacum* leaves (Delannoy *et al.*, 2008). Many cysteine proteases are stored in a stable precursor form in structures called protease precursor vesicles (PPVs) or ricinosomes located within the ER or vacuole (Buchanan-Wollaston, 1997) and these cysteine proteinases are activated upon autophagy making available a ready supply of precursor proteases (Herman and Schmidt, 2004). Many cysteine protease precursors possess vacuolar-targeting sequences within their precursor domains indicating that they are ideally recognized by Golgi-localized vacuolar targeting receptors and that these proteins are directed into the vacuole (Herman and Schmidt, 2004). Generally cysteine proteases of the papain superfamily are initially synthesized by the ER with a large pro-domain which lies in the groove containing the active site precluding proteolytic activity until the enzymes are activated at their destination site (Herman and Schmidt, 2004). Proteases then exit the ER and are transported to and through the Golgi where, depending on the targeting information, they are then transported to the cell's lytic compartment, which in the case of plants is the vacuoles or are secreted into the extracellular space (Herman and Schmidt, 2004). These observations were further justification for not following the secretory pathway route for heterologous protein expression.

In tobacco, cysteine proteases are also involved in the wound response (Ueda *et al.*, 2000). The agro-infiltration process is a form of stress (Rodríguez *et al.*, 2005) to which plants are subjected during the infiltration process which is linked to agrobacterial action as mentioned above. It is likely that many genes encoding the expression of cysteine proteases are up-regulated (Goulet *et al.*, 2010). In plants, wounding causes a shift in the defence protein synthesis pattern localizing resistance to the site of the lesion by enhancing lignification and suberization of the plant cell wall as well as expressing protease inhibitors (McConn *et al.*, 1997; Ebel and Mithofer, 1998). This could potentially negatively affect the production of heterologous proteins in plants depending on the nature of the target protein. In addition to senescence and wounding as contributors to plant stress, the developmental stage and environmental conditions of the plant are also important determinants for proteolytic degradation of heterologous proteins (Stevens *et al.*, 2000). As such, major proteolytic enzymes could jeopardise heterologous protein expression (Stevens *et al.*, 2000). Thus, proteolytic degradation *in planta* remains a primary concern for protein expression and is one of the major challenges that this study aims to address.

1.2.2.7 *Protease inhibitors*

The potential of protease inhibitors as a protective agents to limit degradation of proteins by proteases has been previously demonstrated (Rivard *et al.*, 2006) as cysteine protease inhibitors (cystatins) are tight, reversible inhibitors of papain-like cysteine proteases (Turk and Bode, 1991). The *in vivo* expression of a tomato cathepsin D inhibitor (*S*/CDI) as a protein-stabilizing/protecting agent against aspartic and serine proteases resulted in a 35% - 45% increase in leaf protein content and transient expression of human AACT (α_1 - antichymotrypsin) and was significantly higher in transgenic lines expressing the S/CDI

inhibitor (Goulet *et al.*, 2010). In addition to tomato CDI, bovine aprotinin, a serine protease inhibitor, also enhanced protein stability with regards to potato leaf protein further supporting the idea of protease inhibitors as an "in-built protein stabilizing agent" (Rivard *et al.*, 2006). Further, co-expression of OC-I (oryzacystatin-I) caused a reduction in rubiscase activity through the reduction in cysteine protease activity resulting in a stabilizing effect on Rubisco (Rivard *et al.*, 2006). Van der Vyver *et al.* (2003) also showed that constitutive stable expression of OC-I in tobacco leaves (*N. tabacum* cv. Samsum) increased the amount of soluble proteins in leaves. Reasons for this are still unclear. *N. tabacum* plants were previously transformed with the binary vector pKYOC-I containing the genes encoding for OC-I expression and β-glucuronidase (GUS), each under the regulation of their own CaMV promoter. GUS expression was used as an additional selectable marker with kanamycin resistance for screening for transgenic plant material (Van der Vyver *et al.*, 2003). No detailed study has been carried out yet to demonstrate that co-expression of this cysteine protease inhibitor (OC-I) will increase yield of a transiently expressed heterologous protein such as VP1.

1.2.2.8 *Promoters*

The CaMV 35S promoter from the Cauliflower mosaic virus (CaMV) (Odell *et al.*, 1985) has been the promoter of choice for dicotyledonous plant species for the widespread application in regulating the constitutive expression of foreign genes in genetically engineered plants as well as in transient expression studies using the Agro-infiltration technique. An improvement on the natural 35S CaMV promoter in transcriptional activity is the tandem duplication of 250 bp of upstream sequences which acts as a strong enhancer of promoters (Kay *et al.*, 1987). A modified version of the CaMV 35S promoter, which is tightly repressed by the

Tn*10* encoded Tet repressor in both a transient system and transgenic plants, showed a 500-fold increase in GUS activity upon induction with tetracycline (Gatz *et al.*, 1992). In addition to the 35S promoter, rice promoters *RAamy1A* and *RAamy3D*, induced by sugar starvation, are a good examples of feed-back metabolite repression (Morita *et al.*, 1998; Toyofuku *et al.*, 1998). The ubiquitin *Ubi-1* promoter has been shown to be highly active in monocots and in light of this, a series of vectors high-level expression of selectable and/or screenable marker genes were created to exploit this advantage (Christensen and Quail, 1996) since the CaMV promoter strength is sub-optimal in monocot cells (Christensen *et al.*, 1992). Within the context of our study a CaMV promoter-based system was employed to drive expression of the VP1 gene.

1.3 VP1 prokaryotic and eukaryotic expression systems

1.3.1 Bacterial expression system

In light of the serious threats that FMDV pose to animal health, many approaches have been used to express vaccine candidates in a variety of expression systems in an attempt to curb the spread of the disease. The VP1 protein was first expressed in *E. coli* using a phage λ P_L promoter and plasmid pBR322 (Kupper *et al.*, 1981). The study by Corchero *et al.* (1996) revealed that the *E. coli* BL21 strain lacking the protease La accounts for higher VP1 yields. In contrast, a β-galactosidase-VP1 fusion at the C-terminus resulted in lower yield due to susceptibility to proteases. Historically, the emphasis of heterologous protein expression was placed more on the stability and functionality of the VP1 protein. Recently, yield has become a more important aspect as new expression systems are being explored for vaccine purposes, in particular plant expression systems. Table 1.2 summarises specific studies that have

employed prokaryotic expression systems in order to express regions of VP1 with the intention of creating a vaccine candidate for the disease. Within our study, the *E. coli* strain M15 was used as it permits high-level expression and is easy to handle (QIAexpressionist handbook).

1.3.2 Viral expression system

A VP1 epitope from FMDV serotype O1 encoding amino acids 20-25 representing the FMDV loop, a major immunogenic site, was expressed within the viral small (S) coat protein region of Cowpea mosaic virus (CPMV). The epitope was on the surface of the virion in cowpea plants which was able to elicit an antibody response after expression (Usha *et al.*, 1993). Using a tobacco mosaic virus (TMV)-based vector system, the complete open reading frame coding for VP1 has been expressed in *N. benthamiana* plants using mechanical inoculation (Wigdorovitz *et al.*, 1999b). Table 1.3 summarises specific studies that have employed viral vectors in order to express regions of VP1. Although expression using viral vectors was not employed in this study, their effectiveness and stability as vaccine candidates are well noted based on previous studies.

1.3.3 Plant-based expression systems for VP1 production

The idea of plant-derived vaccines was first conceived in the work produced by Mason *et al.* (1992) where a recombinant Hepatitis B surface antigen (HBsAg) was stably transformed into tobacco plants producing an antigenic version of the surface antigen (Mason *et al.*, 1992). There have been many methods of producing variable regions of the VP1 protein in a variety of plant production systems. These include various transformation methods and host

plants. Currently, the chosen vaccine approach for FMDV is the use of inactivated FMDV viral particles. Although effective for prevention of the disease, production of inactivated virus particles is both costly and extremely risky due to the manipulation of vast amounts of virulent virus which could easily result in virus dissemination (Brown, 1992). Thus there is a need for development of alternative methods for FMDV vaccine production. This can and has been achieved to an extent using bacterial and plants as bioreactors to produce antigens. Table 1.4 summarises studies using transient or stable expression in plants in order to express antigenic regions of the VP1 capsid protein and other studies summarised in Tables 1.2 and 1.3 also highlight the relative success of various strategies.

Table 1.2 Summary of studies conducted on the expression of VP1 antigens in various prokaryotic expression systems.

Epitope/Antigens expressed	FMDV Serotype	Expression type	Expression Host	Size product	Reference
VP1 (1-213 aa)	O1K	MS2 replicase fusion protein	*E. coli*, pBR322	44 kDa	(Kupper *et al.*, 1981)
VP1	O1	LE' fusion[1]	*E. coli*	31 kDa	(Shire *et al.*, 1984)
VP1 (137-162 & 137-213 aa)	O1,BFS	*β*- galactosidase fusion	*E. coli*	120 kDa & 135 kDa	(Broekhuijsen *et al.*, 1986)
VP1 (141-160 aa, 200-207 aa)	A10	Hybrid PhoE fusion proteins	*E. coli*, K12	-	(Agterberg *et al.*, 1990)
VP1 (1-213 aa)	-	Lactose promoter	*E. coli*, BL21	30 kDa	(Neubauer *et al.*, 1992)
VP1	C-S8c1	*β*- galactosidase fusion	*E. coli*	23 kDa	(Corchero *et al.*, 1996)
VP1(134-156 aa)	C-S8c1, serotype C	β-galactosidase fusion	*E. coli strain* MC1061	-	(Feliu and Villaverde, 1998)
VP1	A22, O (V), C, Asia-1, O (B) and Asia-Iz	6x His-tagged fusion	*E. coli* BL21	-	(Suryanarayana *et al.*, 1999)
VP1	A, O, C and Asia 1	6x His-tagged fusion	*E. coli*	-	(Ratish G., 1999)
VP1 (21-40, 135-160, 200-213 aa)	O	GST-fusion	*E. coli*	20 kDa	(Song *et al.*, 2004a)
VP1 (21-40, 135-160, 200-213 aa)	O	6x His-tagged CTB fusion	*Hansenula polymorpha*	20 kDa	(Song *et al.*, 2004b)
VP1 (135-160, 200-213 aa)	O/Taiwan/99	6x His-tagged fusion	*E. coli* Rosetta-2	34 kDa	(Andrianova *et al.*, 2011)

[1] trp leader (L) fused to the last third of the trp E protein gene

Table 1.3 Summary of studies conducted on the expression of VP1 antigens in various viral-based expression systems.

Epitope/Antigens expressed	FMDV Serotype	Expression type	Expression Host	Size product	Immunogenicity	Reference
VP1	O1Campos	Tobacco mosaic virus	*N. benthamiana*	26 kDa	Response - mice	(Wigdorovitz *et al.*, 1999b)
VP1 (142-152, 200-213)	O	Tobacco mosaic virus	*N. tabacum* cv. Samsun	-	Response - guinea pig	(Jiang *et al.*, 2006)
VP1 (141-160)	Serotype O	Cowpea mosaic virus	*Vigna unguiculata* var. blackeye	24 kDa	Reactivity against FMDV antiserum	(Usha *et al.*, 1993)

Table 1.4 Summary of studies conducted on the expression of VP1 antigens in various plant-based expression systems.

Epitope/Antigens expressed	FMDV Serotype	Expression type	Expression Host	Size product	Immunogenicity	Reference
VP1 (135-160)	O1Campos	Stable - *Agrobacterium* mediated	*Arabidopsis thaliana*	30 kDa	Response - mice	(Carrillo *et al.*, 1998)
VP1 (135-160)	O1Campos	Stable - *Agrobacterium* mediated	Alfalfa	30 kDa	Response - mice	(Wigdorovitz *et al.*, 1999a)
VP1 (13-213)	O1Campos	Stable - *Agrobacterium* mediated	*Solanum tuberosum* cv. *Desirée*	30 kDa	Response - mice	(Carrillo *et al.*, 2001)
VP1 (135-160)	O1Campos	Stable - *Agrobacterium* mediated	Alfalfa	26 kDa	Response - mice	(Dus Santos *et al.*, 2002)
Polyprotein P1, protease 3C	O1Campos	Stable - *Agrobacterium* mediated	Alfalfa	-	Response - mice	(Dus Santos *et al.*, 2005)
VP1	Strain O/ China/99	Stable - *Agrobacterium* mediated	Tomato	-	Response - guinea pigs	(Pan *et al.*, 2006)
VP1	-	Stable - tobacco chloroplast transformation	*N. tabacum* L. cv. Large	30 kDa	Reactivity against cow anti-FMDV polyclonal antibodies	(Li *et al.*, 2006)
CTB + VP1 + SEKDEL	-	Stable - *Agrobacterium* mediated	*Solanum tuberosum L.)*	42 kDa	-	(He *et al.*, 2007)
VP1 (135-160)	O1Campos	Chloroplast transformation	*N. tabacum* L. cv. Petite Havana	26 kDa	Response - mice	(Lentz *et al.*, 2010)
VP1(135-160, 200-213), VP4(21-40), 2C(68-76), 3D (1-115), 3D (421-460)	O/Taiwan/99	Transient - *Agrobacterium* mediated	*N. benthamiana*	34 kDa	Response - guinea pigs	(Andrianova *et al.*, 2011)

1.4 Rationale for study

Although inactivated vaccines are effective, their use has been limited by high-containment vaccine production facilities, relatively short-lived immunity that they provide, and the encouragement of a carrier state in some vaccinated animals following contact with FMDV (Su *et al.,* 2007). Thus, novel approaches to vaccine development are emerging with several advantages over conventional vaccine production methods. In this regard, extensive research has been conducted in the past on expressing various different versions and lengths of the VP1 immunogenic epitope from the seven different serotypes of FMDV in a variety of expression systems with the objective to improve expression and ultimately produce an effective vaccine candidate. Key studies have been outlined in the "Introduction" of this dissertation.

In this study the suitability of a plant-based transient expression system using the agro-infiltration technique was compared to an *Escherichia coli* (*E. coli*)-based expression system. A rice-codon optimized version of the VP1 gene from Serotype O, South Korean strain from foot-and mouth disease virus (FMDV) was expressed and any expression of this type of VP1 gene in a plant-based system has so far not been reported. Both *N. benthamiana* and *N. tabacum*, able to perform post-translational modifications, were further used in this study. In particular *N. benthamiana* has been previously used a model plant species for heterologous protein expression (D'Aoust *et al.,* 2009) due to a lower content of secondary compounds interfering in any protein purification process. ER retention was further used in this study to maximise potential expression levels by predisposing the VP1 protein to a cellular environment conducive for folding, post-translational modification, accretion and with a lower protease activity. Studies have shown that many proteases are targeted to the secretory

pathway so there is greater value in retaining proteins within the ER to minimise protein degradation *in planta* as it will most likely be protected from proteases and be expressed in the soluble fraction. A CaMV 35S promoter-based system was employed in this study to drive expression of the VP1 gene as dicotyledonous plant species are used as expression hosts.

N. tabacum was used in this study for testing transient VP1 expression due to the availability of transformed tobacco plants constitutively expressing the rice cysteine protease inhibitor oryzacystatin-I (OC-I) in the cytosol preventing cysteine protease activity in this cellular compartment. These OC-I expressing transformed plants also express β-glucuronidase (GUS) which can be used as a selectable marker for transformation allowing easy screening for transformed plants using either a histochemical or fluorometric assay. Finally, any possible advantage of co-expressing multiple proteins (VP1 and OC-I) rapidly through transient agro-infiltration expression was explored in this study by co-infiltrating VP1 and OC-I in *N. tabacum* plants.

1.5 Research aims and objectives

The research aim of this study was first to compare the efficiency of plant-based (tobacco) transient protein expression system with a bacterial (*E. coli*) expression system and secondly to evaluate the effect of OC-I expression on VP1 protein stability and production in OC-I transformed tobacco.

The study had therefore the following objectives:

1. Express, detect and purify a plant-codon optimized FMD VP1 protein in *E. coli* carrying a His-tag to demonstrate that VP1 can be produced using an *E. coli*-based bacterial expression system.
2. Transiently express the VP1 protein in *N. benthamiana* plants and detect expression with a polyclonal antiserum raised against the FMD virus to demonstrate that VP1 can be transiently expressed in a plant-based system using the agro-infiltration technique.
3. Transiently express the VP1 protein in transgenic *N. tabacum* plants that constitutively express the cysteine protease inhibitor OC-I to evaluate if expression of the inhibitor affects VP1 stability and yield.
4. Transiently co-express simultaneously VP1 and OC-I in *N. tabacum* plants to evaluate if transient co-expression of OC-I with VP1 affects VP1 stability and yield.

2. MATERIALS AND METHODS

2.1 DNA work

2.1.1 Cloning of VP1 coding sequence and sequence analysis

The VP1 gene sequence was kindly provided by Professor Yong Suk Jang (Chonbuk National University, South Korea) and the sequence was cloned into vector pGEM®T-Easy to create plasmid pMYE100. The VP1 coding sequence was already codon-optimized to the codon usage preferred by rice. Plasmid DNA was isolated from transformed *E. coli* cells, strain JM109 [genotype: endA1 glnV44 thi-1 relA1 gyrA96 recA1 mcrB^{+} Δ(lac-proAB) e14- [F' traD36 proAB^{+} lacI^{q} lacZΔM15] hsdR17($r_K^{-}m_K^{+}$)], containing plasmid pMYE100 using the alkaline lysis plasmid mini-prep procedure as described previously (Maniatis *et al.*, 1982). *E. coli* cells were sub-cultured from glycerol stocks onto Luria broth (LB) agar medium supplemented with the antibiotic ampicillin (100 µg mL^{-1}) and incubated O/N at 37°C. Single *E. coli* colonies were cultured in 5 mL of LB medium in McCartney bottles containing LB medium and ampicillin (50 µg mL^{-1}) and were incubated O/N with vigorous shaking at 200 rpm. Plasmid DNA was isolated using the alkaline lysis plasmid mini-prep procedure. A vector map was constructed as shown in Figure 5.1. The presence of plasmid DNA was visually verified on a 1.5% agarose gel (Fig. 3.1a). Plasmids were sequenced using T7 primers (Genotech, South Korea) and sequences were analysed using the Bioedit Sequence Alignment Editor Copyright © 1997-2007 (Tom Hall) and Mega 4 © 1993 – 2003 (K. Tamura, J. Dudley, M. Nei, S. Kumar). The VP1 gene sequence from plasmid pMYE100 was first imported into Mega 4 © 1993 – 2003 where the flanking primer sequences were trimmed. Subsequently, DNA sequence and amino acid sequence information obtained for

the VP1 coding sequence in plasmid pMYE100 were compared to available DNA and amino acid sequence information for the VP1 gene for Serotype O, Pan Asia strain of FMDV on Genbank (Accession number - AF428246) using the alignment tools available in the Bioedit Sequence Alignment Editor Copyright © 1997-2007. Primers for the VP1 coding sequence in plasmid pMYE100 were designed based on sequence information obtained from Genbank for the VP1 gene (Accession number - AF428246) for Serotype O of FMDV.

2.1.2 Primer design, PCR reaction and cloning of PCR product (bacteria)

Two primers were designed containing the restriction enzyme recognition sequences for *Bam*HI at the 5'-end and *Hind*III at the 3'-end. The sequences for each oligo-nucleotide are as follows: forward primer termed VP1(*Bam*HI)-F (Genotech, South Korea) with sequence 5'-GC *GGA TCC* ACC ACC TCC ACA GGT-3' and reverse primer termed VP1(*Hind*III)-R (Genotech, South Korea) with sequence 5'-GC *AAG CTT TTA* CAG AAG CTG TTT CAC-3' which contains the stop codon TAA.

The PCR reaction was set up in a final volume of 20 µL as follows:

1 µL of diluted template DNA (80 ng μL^{-1})

1 µL of *Bam*HI forward primer (10 µM)

1 µL of *Hind*III reverse primer (10 µM)

0.25 µL of Ex Taq polymerase (Takara, Japan)

2 µL of dNTP (Takara, Japan)

2 µL of 10 X Buffer™ (Takara, Japan)

12.75 µL of ddH_2O

The PCR reaction was conducted in a AU/Palm Cycler (Corbett Research Pty Ltd, Australia) according to the following programme:

5 min at 94°C

10 sec at 94°C
30 sec at 55°C
1 min at 72°C
} 30 cycles

7 min at 72°C

A sample of PCR product was analysed on a 1% agarose gel to confirm amplification of the VP1 gene (Fig. 3.1b). Once amplification was confirmed, 15 µL of the PCR sample was gel-purified using the DOKDO Prep™ Gel Extraction Kit (Elpis Biotech, South Korea) and eluted in a final volume of 30 µL using ddH_2O. The eluted DNA was analysed on a 1% agarose gel to confirm the correct VP1 fragment size ensuring that no loss had occurred during the gel purification process.

The purified VP1 amplicon was subsequently cloned into the pGEM®T-Easy vector (Promega, USA) following the manufacturer's instructions. The ligation reaction was set up in a final reaction volume of 10 µL as follows and incubated O/N at 16°C in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA):

1 µL (50 ng) of pGEM®-T-Easy vector (Promega, USA)

3 µL of VP1 purified PCR product (~100 ng μL^{-1})

5 µL of 2 X Rapid Ligation Buffer (Promega, USA)

1 µL of T4 DNA ligase (Promega, USA)

Competent *E. coli* (Top 10) cells (genotype: F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ^-) were prepared and transformed as previously described (Inoue *et al.*, 1990). The ligate (10 µL) was transformed into 50 µL of *E. coli* (Top 10) competent cells as outlined in the Annexure. Putative transformants from LB agar plates [ampicillin (50 µg mL^{-1}), IPTG (1 mM), 5% X-Gal (1 mM)] were cultured O/N in LB medium supplemented with ampicillin (50 µg mL^{-1}) at 37°C at 250 rpm with vigorous shaking and replica plated onto LB agar plates. Plasmid DNA was isolated from colonies using the one-step plasmid mini-prep as described by Chowdury (1991) and analyzed on a 1% agarose gel to confirm the correct VP1 fragment size. Plasmid DNA from putative transformants were then restricted with restriction enzyme *Eco*RI (Beams Biotechnology, South Korea) as a preliminary screening procedure.

The reaction was set up as follows in a final volume of 10 µL consisting of:

2 µL of plasmid DNA (50 ~ 100 ng)

0.5 µL of *Eco*RI restriction enzyme (5 units)

1 µL of 10 X *Eco*RI buffer (DCC Bionet, South Korea)

6.5 µL of ddH_2O

The reaction was incubated at 37°C for 1 hr and heat-inactivated for 20 min at 65°C in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA). A 5 µL sample of the reactions were analysed on a 1% agarose gel to confirm the correct VP1 fragment size (Fig. 3.2). Colonies that did not yield any insert DNA were not selected for further analyses. The alkaline lysis plasmid mini-prep method was used for isolating plasmid DNA from positive colonies for further analyses using restriction enzymes *Bam*HI and *Hind*III. Three reactions were set up as follows. The first single restriction enzyme digest was set up using restriction enzyme *Bam*HI in a final volume of 10 µL containing the following:

1 µL of pure plasmid DNA (50 ~ 100 ng)

0.5 µL of *Bam*HI (5 units, Beams Biotechnology, South Korea)

1 µL of 10 X *Bam*HI buffer (Beams Biotechnology, South Korea)

7.5 µL of ddH_2O

The second single restriction enzyme digest was set up using restriction enzyme *Hind*III and was set up in a final volume of 10 µL in same manner as described for *Bam*HI. The third restriction enzyme digestion was set up with both restriction enzymes in a final volume of 10 µL in the same manner as described for both *Bam*HI and *Hind*III. All three reactions were incubated at 37°C for 1 hr and heat-inactivated at 65°C for 20 min in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA) and analysed on a 1% agarose gel to confirm the correct VP1 fragment size (Fig. 3.3). Positive clones were sequenced. A sequence alignment was done using Bioedit comparing the positive control VP1 coding sequence from the construct pMYE100 with the VP1 coding sequence from the transformant to confirm VP1 gene and amino acid sequence identity. Once the gene and amino acid identities for the VP1 gene sequence in pGEM®T – Easy vector were confirmed, culture from the replica plate for a positive colony was grown O/N at 37°C in LB medium supplemented with ampicillin (50 µg mL^{-1}) with vigorous shaking at 250 rpm. An 800 µL sample of O/N culture was glycerol stocked. The plasmid was designated the name pMYE101. A vector map was constructed as shown in Figure 5.2.

2.1.3 Cloning VP1 coding sequence into plasmid pQE-30

Bacterial expression vector pQE-30 (Qiagen, Valencia, CA) was cultured O/N in LB medium supplemented with ampicillin (100 µg mL^{-1}) and kanamycin (25 µg mL^{-1}). The alkaline lysis

plasmid mini-prep method was used to isolate plasmid DNA for bacterial expression vector pQE-30 (Qiagen, Valencia, CA). See Annexure (Fig. 5.3) for the vector map of bacterial expression vector pQE-30. The isolation of the pQE-30 plasmid DNA was confirmed using a 1% agarose gel. In order to clone the VP1 coding sequence into the plasmid pQE-30, three restriction reactions were set up.

The first single restriction enzyme digestion was conducted using restriction enzyme *Bam*HI and was set up to a final volume of 10 µL containing the following:

1 µL of pure plasmid DNA from construct pMYE101 (~ 100 ng)

0.5 µL of *Bam*HI (5 units, Beams Biotechnology, South Korea)

1 µL of 10 X *Bam*HI buffer (Beams Biotechnology, South Korea)

7.5 µL of ddH_2O

The second restriction enzyme digest was done using restriction enzyme *Hind*III (Bioneer, South Korea) and was set up to a final volume of 10 µL in the same manner as described for *Bam*HI. The third double restriction enzyme digestion was conducted using both restriction enzymes *Bam*HI and *Hind*III and was set up to a final volume of 30 µL containing the following:

10 µL of pure plasmid DNA from construct pMYE101 (~ 1 µg)

1 µL each of *Bam*HI and *Hind*III (10 units)

3 µL of 10 X Buffer B (Beams Biotechnology, South Korea)

15 µL of ddH_2O

All three reactions were incubated at 37°C for 2 hr and heat-inactivated at 65°C for 20 min in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA) and analysed on a 1%

agarose gel. The VP1 fragment was subsequently gel purified using the DOKDO Prep™ Gel Extraction Kit (Elpis Biotech, South Korea) and eluted in a final volume of 30 µL in ddH_2O.

Similarly, the bacterial expression vector pQE-30 was also restricted with restriction enzymes *Bam*HI and *Hind*III in the same manner as plasmid pMYE101. The third double restriction enzyme digestion was set up using *Bam*HI first and was set up in a final volume of 30 µL containing the following constituents:

14.5 µL of pure plasmid DNA (~ 1.4 µg)

1 µL of *Bam*HI (10 units)

3 µL of 10 X Buffer B (Beams Biotechnology, South Korea)

11.5 µL of ddH_2O

This reaction was incubated at 37°C for 2 hrs in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA). A 5 µL sample of the third restriction enzyme reaction was analysed using agarose gel electrophoresis to ascertain linearization of plasmid DNA with the *Bam*HI restriction enzyme after a 2 hrs incubation period at 37°C. Once linearization was confirmed, the following constituents were added to the third restriction enzyme reaction: 1 µL of *Hind*III was added, 3 µL of 10 X buffer B and finally 1 µL of ddH_2O taking the final volume back up to 30 µL and further incubated at 37°C for 2 hr and heat-inactivated at 65°C for 20 min in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA). The samples were analysed on a 1% agarose gel. The linearised 3.5 kb pQE-30 fragment was gel-purified using the DOKDO Prep™ Gel Extraction Kit (Elpis Biotech, South Korea) and eluted in a final volume of 30 µL in ddH_2O. Both the VP1 insert DNA and pQE-30 vector DNA were analysed on a 1% agarose gel prior to ligation to confirm ratios.

The ligation reaction was set up in a final volume of 10 µL as follows:

2 µL of gel purified vector DNA (pQE-30)

6 µL of gel purified insert DNA (VP1)

1 µL of ligase buffer (Takara Bio Inc., Japan)

1 µL of T4 DNA ligase (Takara Bio Inc., Japan)

The reaction was incubated O/N at 16°C in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA). The pGEM®T – Easy protocol for transformation was followed in order to transform the pQE-VP1 ligate into competent *E. coli* (Top 10) cells. The one-step plasmid mini-prep method was used to isolate plasmid DNA from putative transformants containing the pQE-VP1 construct from *E. coli* (Top 10) cells. The plasmid DNA was analysed on a 1% agarose gel. Putative transformants were replica plated onto LB plates supplemented with ampicillin (50 µg mL^{-1}) and sub-cultured in LB medium supplemented with ampicillin (50 µg mL^{-1}). Plasmid DNA was isolated from putative transformants and restriction enzyme digestions were conducted in same manner as described above for *Bam*HI and *Hind*III to confirm the presence of the VP1 sequence insert. Once the presence of the VP1 sequence was confirmed in bacterial expression vector pQE-30 in *E. coli* (Fig. 3.4), a single colony from the replica plate was grown O/N at 37°C in LB medium supplemented with ampicillin (50 µg mL^{-1}) and glycerol stocked. The plasmid was named pMYE103. See Annexure (Fig. 5.5) for the vector map for pMYE103.

The alkaline lysis plasmid DNA mini-prep protocol was used to isolate plasmid DNA from pMYE103 for transformation into *E. coli* bacterial expression host strain M15. See manufacturers guidelines in the QIAexpressionist's handbook (Qiagen, Valencia, CA) for the methodology of preparation of competent *E. coli* M15 (genotype: $NaI^{S,}$ Str^{S}, Rif^{S}, Thi^{-}, Lac^{-},

Ara^{+}, Gal^{+}, Mtl^{-}, F^{-}, RecA^{+}, Uvr^{+}, Lon^{+})] cells and the transformation of pMYE103 and control construct pQE-30 into *E. coli* M15 cells. Putative pQE-VP1-M15 and control pQE-30-M15 transformants were selected and replica plated on LB agar plates supplemented with kanamycin (25 µg mL^{-1}) and ampicillin (100 µg mL^{-1}) and inoculated in LB medium supplemented with kanamycin (25 µg mL^{-1}) and ampicillin (100 µg mL^{-1}) and incubated O/N at 37°C with vigorous shaking at 200 rpm. The one step plasmid mini-prep technique was used to isolate plasmid DNA from putative transformants which were analysed on a 1% agarose gel. Three colonies were selected for further restriction enzyme digestion analysis using *Bam*HI and *Hind*III as described previously. Once the presence of the VP1 sequence in expression vector pQE-30 was confirmed (Fig. 3.5), a single colony from the replica plate was grown O/N at 37°C in LB medium supplemented with kanamycin (25 µg mL^{-1}) and ampicillin (100 µg mL^{-1}) and glycerol stocked. The corresponding empty control plasmid vector named pMYE105 was also transformed into *E. coli* M15 cells.

2.1.4 Cloning VP1 coding sequence into binary vector pMYV497 (plants)

Two primers were designed based on sequence information from NCBI on the VP1 gene (Accession number - AF428246) from FMDV, serotype O containing restriction enzyme recognition sequences for *Bam*HI at the 5'-end and *Kpn*I at the 3'-end. The sequences for each oligo-nucleotide are as follows with forward primer termed VP1(*Bam*HI)-F (Genotech, South Korea) with sequence 5'-GC *GGA TCC* ACC ACC TCC ACA GGT-3' and for reverse primer VP1(*Kpn*I)-R (Genotech, South Korea) with sequence 5'-GC *GGT ACC* CAG AAG CTG TTT CAC A-3'. The PCR reaction, PCR program and PCR purification was carried out as described above using primers VP1(*Bam*HI) and VP1(*Kpn*I)-R. The purified VP1 amplicon (Fig. 3.6) was cloned into pGEM®T-Easy vector (Promega, USA) following the

manufacturer's instructions. The ligation reaction between the VP1 fragment and pGEM®T-Easy vector was carried out as described above. The transformation protocol was followed according to the Promega Technical Manual for pGEM®T and pGEM®T Easy Vector Systems. Putative transformants were selected from LB plates supplemented with kanamycin (50 µg mL^{-1}) and were cultured O/N in LB medium supplemented with kanamycin (50 µg mL^{-1}) at 37°C at 250 rpm with vigorous shaking. Colonies were replica plated on LB plates supplemented with kanamycin (50 µg mL^{-1}). Plasmid DNA from putative transformants was isolated using the one-step mini-prep method. Plasmid DNA from putative transformants were restricted with *Eco*RI (Beams Biotechnology, South Korea) as described above. A 10 µL sample from all the restriction reactions were analysed on a 1% agarose gel as shown in Figure 3.7. The alkaline lysis plasmid mini-prep method was used for isolating plasmid DNA from a positive colony for further analyses using restriction enzymes *Bam*HI and *Kpn*I. Three reactions were set up. The first single restriction enzyme digestion was done using *Bam*HI (Beams Biotechnology, South Korea) and was set up to a final volume of 10 µL containing the following:

2 µL of pure plasmid DNA (50 ~ 100 ng)

0.5 µL of restriction enzyme *Bam*HI (5 units)

1 µL of 10 X *Bam*HI buffer (Beams Biotechnology, South Korea)

6.5µL of ddH_2O

The second single restriction enzyme digestion was set up using restriction enzyme *Kpn*I (Beams Biotechnology, South Korea) and was set up to a final volume of 10 µL in the same manner as described above for the *Bam*HI restriction digest. The third double restriction enzyme digestion was set up using both restriction enzymes *Bam*HI and *Kpn*I and was set up to a final volume of 10 µL in the same manner as described above for *Bam*HI and *Kpn*I. All

three reactions were incubated at 37°C for 1 hr and then heat-inactivated at 65°C for 20 min in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA) and then analysed on a 1% agarose gel as shown in Figure 3.8. This clone was subsequently sequenced. A sequence alignment was done using Bioedit comparing the positive control VP1 gene sequence from plasmid pMYE100 with the VP1 gene sequence from the transformant to confirm VP1 gene and amino acid sequence identity. Once identities for the VP1 gene sequence in pGEM®T – Easy vector were confirmed, culture from the replica plate for this clone was grown O/N at 37°C in LB medium supplemented with kanamycin (50 µg mL^{-1}) with vigorous shaking at 200 rpm and glycerol stocked. The plasmid was designated the name, pMYE102. See annexure for vector map of construct pMYE102 in Figure 5.4.

The alkaline lysis plasmid DNA mini-prep was used to isolate plasmid DNA for binary vector pMYV497 and plasmid pMYE102. See Figure 5.6 in annexure for the vector map of plasmid pMYV497. Restriction enzyme digestion was conducted on plasmid DNA from plasmid pMYE102 using *Bam*HI, *Kpn*I and both enzymes as described above and then analysed on a 1% agarose gel. Once the correct size for the VP1 sequence was confirmed, the VP1 fragment was gel purified using the DOKDO Prep™ Gel Extraction Kit (Elpis Biotech, South Korea) and eluted in a final volume of 30 µL of ddH_2O.

The binary vector pMYV497 was linearized using restriction enzymes *Bam*HI and *Kpn*I (Beams Biotechnology, South Korea) in a final volume of 30 µL containing the following:

15 µL of pure plasmid DNA (500 ~ 1 µg)

1 µL each of restriction enzymes *Bam*HI and *Kpn*I (10 units)

3 µL of 10 X Buffer A (Beams Biotechnology, South Korea)

10 µL of ddH_2O

The reaction was incubated at 37°C for 3 hr and heat-inactivated at 65°C for 20 min in an Echotherm chilling/heating plate (Torrey Pines Scientific, USA) and then analysed on a 1% agarose gel. Once linearization of the vector was confirmed, the vector fragment was gel purified using the DOKDO Prep™ Gel Extraction Kit (Elpis Biotech, South Korea) and eluted in a final volume of 30 µL of ddH_2O. The insert and vector DNA were analyzed on a 1% agarose gel to confirm their presence after gel purification before ligation.

The ligation reaction between vector pMYV497 and VP1 insert was set up in a final volume of 10 µL with the following constituents and incubated O/N at 16°C:

4 µL of pMYV497 vector DNA

4 µL of VP1 insert DNA

1 µL of 10 X Ligase buffer (Takara, Japan)

1 µL of T4 DNA ligase (Takara, Japan)

The transformation into *E. coli* (Top 10) cells was followed according to the Promega Technical Manual for pGEM®T and pGEM®T Easy Vector Systems. The transformation mix was plated onto LB plates supplemented with kanamycin (50 µg mL^{-1}) and incubated O/N in a 37°C incubator. The alkaline lysis plasmid DNA method was used to isolate plasmid DNA from putative transformants. A single restriction enzyme digestion was carried out on the

putative transformants using both restriction enzymes *Bam*HI and *Kpn*I as described above and analysed on a 1% agarose gel as shown in Figure 3.9. Once the presence of the VP1 sequence in pMYV497 was confirmed a single colony from the replica plate was grown O/N at 37°C in LB medium supplemented with kanamycin (50 µg mL^{-1}) and glycerol stocked. The plasmid was designated the name pMYE108. See Annexure (Fig. 5.7) for the vector map of plasmid pMYE108.

Agrobacterium tumefaciens LBA4404 was transformed with plasmid pMYE108 using the tri-parental mating method as previously described (Wang, 2006). Plasmid DNA from putative transformants was isolated using the alkaline lysis method and analysed on a 1% agarose gel. Colonies were replica plated on an LB plate supplemented with rifampicin (100 µg mL^{-1}) and kanamycin (50 µg mL^{-1}). Plasmid DNA from a single colony was re-transformed into competent *E. coli* (Top 10) cells. The transformation protocol was followed according to the Promega Technical Manual for pGEM®T and pGEM®T Easy Vector Systems. The transformation mix was plated onto LB plates supplemented with kanamycin (50 µg mL^{-1}) and incubated O/N in a 37°C incubator. The alkaline lysis method was used to isolate plasmid DNA from putative transformants. Restriction enzyme digestion analysis was conducted using *Bam*HI, *Kpn*I and both enzymes as described above and were analysed on a 1% agarose gel as shown in Figure 3.10. Once the presence of the VP1 sequence in vector pMYV497 was confirmed the clone from the replica plate was grown O/N at 37°C in LB medium supplemented with kanamycin (50 µg mL^{-1}) and rifampicin (100 µg mL^{-1}) and glycerol stocked. For a summary of all the constructs created within the study please refer to Figure 5.8 within the Annexure under the Plasmid Maps section.

2.2 Plant material - tobacco growth

Six week old *N. benthamiana* plants were grown in a tissue culture growth chamber and maintained at 26°C/20°C day/night temperature cycle and a 12/12 hours light/dark cycle. Plants were obtained from Professor Moon Sik Yang (Chonbuk National University, South Korea). Photosynthesis photon flux density during light phase was 2000 ± 50 µmol m^{-2} s^{-1}. Plants were grown at a relative humidity of 60%. Eight week old *N. tabacum* L. cv. Samsum plants were grown in a greenhouse and maintained at 26°C/ 20°C day/night temperature cycle and a 12/12 hours light/dark cycle. Transformed and non-transformed *N. tabacum* L. cv. Samsum plants after self-fertilizing of T4/5 plants were obtained from Professor Karl Kunert (FABI, Plant Science Department, University of Pretoria, 0002, South Africa). Tobacco were transformed with the construct pKYOC-I (Fig. 5.8c) allowing constitutive expression of OC-I and GUS under the control of a double 35S CaMV (cauliflower mosaic virus) promoter (P70) and a single 35S promoter sequence, respectively. GUS served as a selectable marker (Van der Vyver *et al.*, 2003) for selecting transformants. All plants were grown in the greenhouse with a photosynthesis photon flux density during light phase of 600 ± 50 µmol m^{-2} s^{-1} and a relative humidity of 60%.

2.3 Agro-infiltration

Agrobacterium-mediated transient expression system was used to infiltrate the tobacco leaf surface using the syringe agro-infiltration technique (D'Aoust *et al.*, 2009). The *Agrobacterium tumefaciens* strain LBA 4404 containing the VP1 coding sequence in the binary vector pMYE108 (Fig. 5.7) was streaked out onto LB (Luria-broth) [1% bacto-tryptone (w/v), 0.5% bacto-yeast extract (w/v), 1% sodium chloride (w/v) and 1.2% bacto-

agar (w/v), pH 7.4] agar plates supplemented with rifampicin (100 µg mL^{-1}) and kanamycin (50 µg mL^{-1}) and grown for two days at 28°C. Single colonies were inoculated into 5 mL of LB medium (1% bacto-tryptone (w/v), 0.5% bacto-yeast extract (w/v), 1% sodium chloride (w/v) pH 7.4) supplemented with the above-mentioned antibiotics and grown at 28°C with vigorous shaking at 200 rpm for two days. Fifty millilitres of modified yep medium (10 g L^{-1} bacto-peptone, 10 g L^{-1} yeast extract, 5 g L^{-1} NaCl, 10 mM MES, 10 mM $MgCl_2$, pH 5.5) containing the above mentioned antibiotics and acetosyringone (20 µM) were inoculated with 5 mL of two day-old *Agrobacterium* culture and grown at 28°C O/N with vigorous shaking at 200 rpm. The O/N culture was spun down in a large centrifuge (Beckman – Avanti J-25 large centrifuge) at 5000 x g for 15 min at 4°C. The supernatants were discarded and the pellets were re-suspended with modified MSO medium [1.32 g of Murashige & Skoog (Duchefa Biochemie), 10 mM MES, 10 mM $MgCl_2$, pH 5.6] and spun down at 5000 x g for 15 min at 4°C. After centrifugation, the pellets were re-suspended in 30 mL of modified MSO medium and 5 mL of that suspension was placed in a new falcon tube and diluted using modified MSO in order to obtain a final OD_{600} (optical density at a wavelength of 600 nm) of 1. Acetosyringone (200 µM) was added to the suspension which was incubated for 2-3 hr at room temperature (25°C). The infiltrated *N. benthamiana* leaves were washed in ddH_2O (double distilled water) and placed on Whatman filter paper (Whatman International Ltd., England) before storing on MSO (10 mM MES, 10 mM $MgCl_2$, pH 5.6) agar plates. On average, five leaves per plant were infiltrated. Infiltrated leaves were incubated in the tissue culture incubator for the time-course duration of three, five and seven days post-infiltration. The syringe agro-infiltration procedure outlined above was also used to transiently express VP1 in transgenic and control *N. tabacum* L. cv. Samsun plants. Leaves were left intact on the plant and incubated for the time-course duration of five and seven days post-infiltration to simulate senescent conditions. Plants were divided into two groups: OC-I expressing and non

OC-I expressing based on both fluorometric GUS activity detection and cysteine proteinase inhibitor immuno-detection carried out prior to transient agro-infiltration. The data are summarized in Table 3.3. Selected plants were infiltrated only with the VP1 binary vector and putative VP1 content was determined using immune-detection. An experiment was also carried out where an OC-I and non-OC-I expressing plant were co-infiltrated with both the pKYOC-I binary vector and the VP1 expressing binary vector (pMYE108) in order to investigate the effect of transient OC-I co-expression. The phenotype of agro-infiltrated OC-I tobacco leaves is shown in Figure 3.16.

2.4 Protein work

2.4.1 Protein expression (bacteria)

2.4.1.1 *Cell growth, protein expression and isolation*

Single colonies harbouring plasmid pMYE103 and empty vector control plasmid pMYE105 were inoculated into 5 mL of LB medium supplemented with ampicillin (100 µg mL^{-1}) and kanamycin (25 µg mL^{-1}) and grown O/N at 37°C with vigorous shaking at 200 rpm. A sample (1 mL) of O/N culture from pMYE103 and control plasmid pMYE105 was spun down at 15 000 x g for 1 min in a table-top centrifuge (Eppendorf centrifuge 5415 R, Germany) and stored at -20°C freezer serving as non-induced controls.

Two separate flasks containing 10 mL of pre-warmed LB medium supplemented with ampicillin (100 µg mL^{-1}) and kanamycin (25 µg mL^{-1}) were inoculated with 500 µL of O/N culture from plasmid pMYE103 and control plasmid pMYE105 respectively and were grown

at 37°C with vigorous shaking at 200 rpm until the OD_{600} reached 0.5 – 0.7. The OD_{600} was checked with a spectrophotometer (Biospec – 1601, Shimadzu, USA) before IPTG induction. VP1 expression was induced by adding IPTG to a final concentration of 1 mM. Cells were grown for an additional 4 – 5 hrs and then harvested by centrifugation for 1 min at 15000 x g. Supernatants were discarded and cell pellets were collected and stored at –20°C until analysed by SDS-PAGE and immuno-blotting.

2.4.1.2 *Protein purification*

For purification, the above process was scaled up to a 200 ml volume flask used for induction of VP1 protein expression as opposed to using 10 ml. Cell pellets of non-induced and IPTG-induced cells were re-suspended in 10 mL of Buffer Z (8M Urea, 100 mM NaCl, 20 mM Hepes, pH 8.0) in falcon tubes. Cells were sonicated for 20 min on ice and left in ice for 5 min (frothing was avoided during the sonication process). The solution became a clear translucent colour when sonication was effectively completed. The solution was centrifuged in falcon tubes at 5000 rpm for 15 min at 4°C. The supernatants were transferred to fresh falcon tubes. A sample (1 mL) of the supernatant was stored at 4°C. Imidazole (1 M) (100 µL) was added to 10 mL of supernatant giving a final imidazole concentration of 10 mM. This was termed the lysis sample. A purification column (Bio-rad, Korea) was loaded with 700 µL of Ni-NTA resin (Qiagen, Valencia, CA). The column was loaded with 1 mL of ddH_2O which flowed through the Ni-NTA resin and out under the column. Sterile petri-dishes were used to catch fractions. The column was then loaded with 1 mL of 10 mM imidazole in Buffer Z. The lysis sample was loaded onto the column, 5 mL at a time. A sample (1 mL) of the flow-through was kept and stored at 4°C. This flow-through was reloaded to increase yields. The sample was washed three-times with 10 mL of 10 mM

imidazole in buffer Z. A sample (1 mL) of the wash solution was collected and stored at 4°C. The protein was eluted using 250 mM imidazole in Buffer Z. The elution was done three-times. A sample (1 mL) of elution one, 500 µL sample of elution two, 500 µL sample of elution three were collected and stored at 4°C. All fractions were then analyzed using 12% SDS-PAGE. The experiment was conducted three-times and protein concentrations obtained for Elution 1 and 2 are summarized in Table 3.2. The separated protein bands were analysed using SDS-PAGE (Figs. 3.11a, 3.12a and 3.13a) and transferred from the gel to Hybond C membrane (Promega, USA) to carry out Western blot analyses (Figs. 3.11b, 3.12b and 3.13b) as described under Western blotting.

2.4.2 Protein extraction (plants)

Leaves were crushed using a mortar, pestle and liquid nitrogen. Equal amounts of leaf material were placed into 1.5 mL microfuge tubes and Arakawa buffer was added in a 1:1 ratio, 1 being the weight of the sample in milligrams and 1 being the amount of Arakawa buffer (200 mM Tris-Cl, pH 8.0, 100 mM NaCl, 400 mM sucrose, 10 mM EDTA, 14 mM 2-mercaptoethanol, 1 mM phenylmethylsulfonylfluoride (PMSF), 0.05% Tween-20) in microlitres. Leaves were then crushed with a drill homogenizer. The buffer-leaf mixture was centrifuged for 20 min at 20 000 x g. The supernatant containing the soluble protein extract was pipetted into a new 1.5 mL microfuge tube and stored at - 80°C until analysed.

2.4.3 Protein determination

The TSP concentration for *N. benthamiana* and *N. tabacum* protein extracts was determined using the Bradford method using a commercially available protein determination kit

following the instructions provided by the supplier (Sigma Bradford Reagent B6916, Sigma®-Aldrich, Germany). Optical densities for each sample were measured in a spectrophotometer at 595 nm and protein concentrations were extrapolated using a BSA standard curve. A BSA Standard curve was set up from a 10 mg/ml BSA stock solution. The OD of each sample with a different amount of BSA (0, 0.25, 0.5, 1 and 1.4 mg ml^{-1}) was measured at a wavelength of 595 nm in duplicate in a spectrophotometer. A standard curve was constructed and the R^2 value determined and only a R^2 value of at least 0.9 was accepted for subsequent use in determining protein concentrations.

2.4.4 Protein detection

2.4.4.1 *Immuno-detection of VP1 (bacteria)*

Bacterial cell pellets were re-suspended in 2 ml PBS buffer (8 g NaCl, 0.2 g KCl, 2.7 g $Na_2HPO_4.7H_2O$, 0.24 g KH_2PO_4, pH 7.5) and were sonicated for 10 sec with a Branson sonicator (Wolf Labs, UK) and then centrifuged at 13000 rpm in an Eppendorf centrifuge at 4°C for 30 min. Resulting supernatants were carefully sucked off and placed into a fresh 1.5 mL microfuge tube. The supernatant (50 µL) sample was boiled with 5x SDS-containing reducing sample buffer (10% w/v SDS; 10 mM β-mercapto-ethanol; 20% v/v glycerol; 0.2 M Tris-Cl, pH 6.8; 0.05% w/v bromophenol blue) for 5 min in a water bath at 95°C and then cooled on ice for 5 min. A 12% sodium dodecylsulfate polyacrylamide gel (SDS-PAGE) was set up following the method outlined previously (Laemmli, 1970). After SDS-PAGE, separated protein bands were transferred to a Hybond C membrane (Promega, USA) using a Mini Trans Blot electrophoretic transfer cell (Bio-Rad, USA) for 2 hrs at 130 mA in transfer buffer (50 mM Tris, 40 mM glycine, 20% methanol, pH 8.3) at 45V at 4°C. Nonspecific

antibody binding was blocked by incubating the membrane in 20 ml of 5% skim milk (Difco™ Laboratories, Becton, Dickinson and Company) O/N in PBST buffer (8 g NaCl, 0.2 g KCl, 2.7 g $Na_2HPO_4.7H_2O$, 0.24 g KH_2PO_4, 0.05 % Tween-20, pH 7.5). This was followed by three washes of membrane in PBST buffer for 5 min. The membrane was then incubated for 3 hrs in a 1:5000 dilution of an anti-His antiserum raised in mice (Qiagen, Valencia, CA) in PBST buffer followed by three washes of the membrane in a PBST buffer for 5 min. The membrane was incubated for 1 hr in a 1:7000 dilution of a goat anti-mouse IgG conjugated to alkaline phosphatase (Promega S3731, Madison, WI). The membrane was washed again three-times in PBST buffer for 5 min before development. Membranes were developed with nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) (Sigma®-Aldrich, Germany) for 20 min to visualize alkaline phosphatase activity. An identical procedure was followed to ascertain the immunogenicity of the *E. coli* derived VP1 protein using an Anti-FMDV polyclonal antiserum (1:7000 dilution) raised in rabbit followed by hybridization with a goat anti-rabbit IgG conjugated to alkaline phosphatase. The anti-FMDV polyclonal antiserum was kindly provided by Dr Francois Maree (Ondersterpoort Veterinary campus, University of Pretoria, South Africa). Membranes were developed with BCIP/NBT solution.

2.4.4.2 *Immuno-detection of VP1 (plants)*

TSP was extracted from *N. benthamiana* and *N. tabacum* plants using a protein extraction buffer containing PMSF which is a commercial inhibitor generally used as an inhibitor of serine protease activity, but can also act against certain cysteine proteases, particularly those in the papain family (Alonso *et al.*, 1996). TSP from either *N. benthamiana* or *N. tabacum* (30 µg of total soluble protein) were analysed using SDS-PAGE and immuno-blotting. TBST

buffer (12.1 g Tris, 9 g NaCl, 0.1% Tween-20, pH 7.5) was used in place of PBST buffer for O/N blocking of plant protein samples. The VP1 protein produced in *E. coli* was used as a positive control in the immuno-blotting procedure using a 1:7000 dilution of the FMDV antiserum and a 1:7000 dilution of goat anti-rabbit IgG conjugated with alkaline phosphatase (Promega S3731, Madison, WI). Membranes were washed three times after incubation with the primary antibody and after incubation with the secondary antibody membranes were washed twice in TBST buffer and once in TMN buffer (100 mM Tris, 5 mM $MgCl_2.6H_2O$, 100 mM NaCl, pH 9.5) before development. Membranes were developed with BCIP/NBT solution. Approximate values for the VP1 protein were quantified using quantitative densitometry (Abramoff *et al.*, 2004).

2.4.4.3 *Immuno-detection of OC-I (plants)*

A 15 % SDS-PAGE was set up according to the method of Laemmli (1970) and TSP from transformed *N. tabacum* plants (30 µg of total soluble protein) were analyzed for the OC-I expression using immuno-blotting (Laemmli, 1970). In order to detect OC-I expression a 1:5000 dilution of the OC-I antiserum raised in rabbit was used and a 1:10000 dilution of goat anti-rabbit IgG conjugated with alkaline phosphatase (Promega S3731, Madison, WI). Membranes were developed with BCIP/NBT solution.

2.4.4.4 *Quantitative densitometry analysis*

Approximate values for the VP1 protein were quantified using Image J software (Abramoff *et al.*, 2004). Purified *E. coli*-derived VP1 protein was used as a standard with which to compare and quantify the plant-derived VP1 protein. Using the relative density function

within Image J, estimates of expressed VP1 protein values were obtained and are expressed as percent of total soluble protein.

2.4.5 Enzyme activity assays

2.4.5.1 *Gus activity*

The MUG (4-methyl umbelliferyl glucuronide) assay was used to determine β-glucuronidase (GUS) activity in plant extracts. A tobacco leaf extract (50 µL containing 20 µg of total soluble protein) in Arakawa extraction buffer was added to 200 µL of a MUG buffer [50 mM $NaPO_4$ (pH 7.0); 0.1% Triton X-100; 10 mM Na_2EDTA; 0.1% sodium lauryl sarkosine; 10 mM β-mercaptoethanol and 1 mM MUG]. GUS activity was measured by monitoring the release of fluorescent MU (methyl umbelliferyl) at time 0, 30 and 60 min. Reactions were incubated at 37°C and for measuring release of MU at each time point, 50 µL of reaction mixture was added to 950 µL 0.2 M Na_2CO_3 stopping buffer. MU fluorescence was measured using a fluorometer (BMG FluoStar Galaxy). A MU standard curve was constructed using a 1 mg ml^{-1} MU solution stock. MU stock (10 µL) was added to 990 µL of extraction buffer yielding a concentration of 10 µg ml^{-1} MU. Two-fold dilutions were prepared serially yielding concentrations of 5, 2.5, 1.25, 0.625 and 0.3125 µg ml^{-1} MU. Tobacco leaf protein extracts (20 - 40 µg total soluble protein) from agro-infiltrated *N. tabacum* plants were initially screened by measuring GUS activity. Measurements of activities in control (non-OC-I expressing) and experimental (OC-I expressing) plants were taken at day 0, day 5 and day 7 post infiltration. All tobacco foliar protein extracts were diluted to the same concentration using Arakawa buffer before measuring the fluorescence.

2.4.5.2 *Cysteine protease activity*

Total soluble protein from foliar extracts (60 µg total soluble protein) of *N. tabacum* plants were used for measuring cysteine protease activity (cathepsin L-like activity) in extracts. Activity was measured in a 50 mM sodium phosphate buffer (pH 6.0) containing 10 mM L-cysteine (Sigma®-Aldrich, Germany). Each protein sample (20 µL) containing 60 µg total soluble protein was mixed with sodium phosphate buffer (72 µL) and each sample was individually transferred into black, flat-bottom polysorp 96 well plates (Nunc, AEC Amersham) for measuring fluorescence. Fluorescence development was measured with a fluorometer (BMG FluoStar Galaxy) at 25°C with excitation and emission wavelengths of 360 nm and 450 nm, respectively. After monitoring initial fluorescence, 8 µL of the cathepsin L substrate Z-Phe-Arg-7-amido-4-methylcoumarin hydrochloride (Z-Phe-Arg-MCA) was added to each well to start the enzymatic reaction which was measured kinetically over a 10 min time period. In the reaction, fluorescent α-amino 4-methylcoumarin (MCA) was released when a cysteine protease hydrolyzes Z-Phe-Arg-MCA.

2.5 Statistical analysis

Statistical analysis was carried out using a Microsoft Exel software 2007 version 12 (Microsoft Corporation). Comparison between OC-I expressing and non-OC-I expressing groups were made using student's t-test. The results of the study are expressed as means and standard error range. A value of $P < 0.05$ was considered as statistically significant.

2.6 Bioinformatic analysis

Molecular graphics and analyses were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health (National Center for Research Resources grant 2P41RR001081, National Institute of General Medical Sciences grant 9P41GM103311).

The N-glycosylation status of the VP1 protein was predicted using NetNGly available at http://www.cbs.dtu.dk/services/NetNGlyc/. By default, predictions were done only on the Asn-Xaa-Ser/Thr sequences (including Asn-Pro-Ser/Thr).

PROTEIN CALCULATOR v3.3 by Chris Putnam (available at (http://www.scripps.edu/~cdputnam/protcalc.html) was used to ascertain the molecular weight of the VP1 protein.

The online PeptideCutter (http://web.expasy.org/peptide_cutter/ Expasy, Bioinformatics Resource Portal, Swiss Institute of Bioinformatics) was used to predict potential cleavage sites cleaved by proteases or chemicals in the VP1 protein sequence. "PeptideCutter predicts potential cleavage sites cleaved by proteases or chemicals in a given protein sequence. PeptideCutter returns the query sequence with the possible cleavage sites mapped on it and /or a table of cleavage site positions." The parameter of all enzymes and chemicals was selected before submitting the query.

3. RESULTS

3.1 DNA work

3.1.1 Cloning of VP1 coding sequence and sequence analysis

Plasmid DNA for plasmid pMYE100 containing the VP1 gene was isolated (Fig. 3.1A). The VP1 gene from plasmid pMYE100 was sequenced to confirm VP1 identity (see Annexure for the sequence alignment between the VP1 sequence (Genbank) and the codon-optimized VP1 sequence from pMYE100). The plasmid pMYE100 was sequenced and the identities between the VP1 sequence from Genbank and that obtained from plasmid pMYE100 were compared revealing a 0.8794992 or 87.9% identity as determined by the "Calculate identity/similarity for two sequences" function available in Bioedit Sequence Alignment Editor Copyright © 1997-2007 (Tom Hall). Lower than 100% identity was due to codon optimization of VP1 (see Annexure for the amino acid sequence alignment). The identities between the VP1 amino sequence from Genbank and that obtained from plasmid pMYE100 was 1.0000000 or 100% as determined by the "Calculate identity/similarity for two sequences" function available in Bioedit Sequence Alignment Editor Copyright © 1997-2007 (Tom Hall). Once the gene identity was confirmed the VP1 sequence was subsequently PCR amplified with primers VP1(*Bam*HI)-F and VP1(*Hind*III)-R obtaining a VP1 amplicon with a size of ~ 650 bp (Fig. 3.1B).

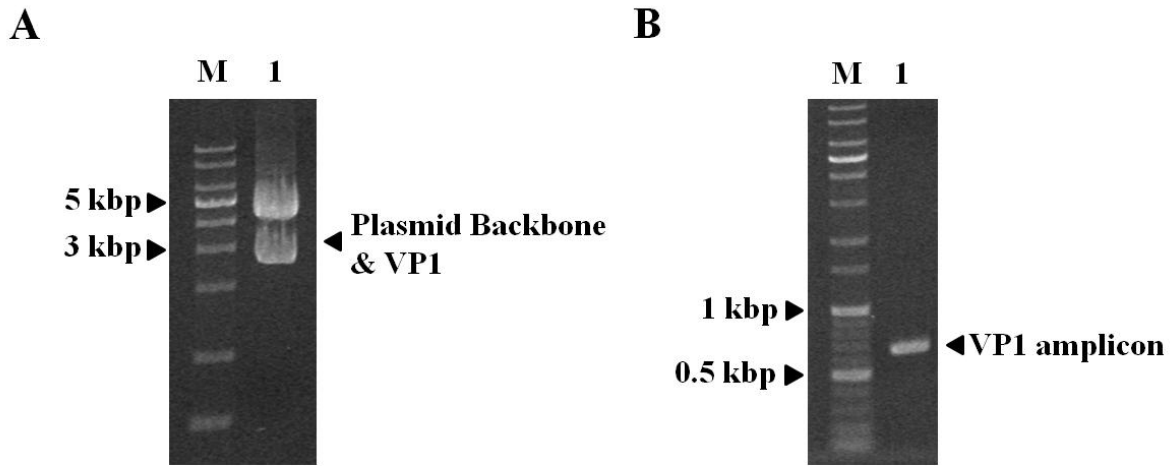


Figure 3.1 Agarose gel electrophoresis of plasmid pMYE100 and PCR amplification of VP1 coding sequence with primers VP1(*Bam*HI)-F and VP1(*Hind*III)-R from plasmid pMYE100 template DNA. (A) M represents the 1 kb DNA ladder. Lane 1 represents undigested pMYE100 plasmid DNA containing the VP1 gene in the vector pGEM®-T Easy. (B) M represents a 1 kb molecular DNA plus ladder and lane 1, the VP1 amplicon PCR amplified from pMYE100 plasmid template DNA with a size of ~ 650 bp.

Subsequently, VP1 was sub-cloned back into the pGEM®-T-Easy vector. VP1 sequence identity was confirmed by restriction enzyme analysis. Putative pGEM®-T-VP1 transformants were initially restricted with *Eco*RI (Fig. 3.2) to confirm the presence of the VP1 sequence. *Eco*RI restriction analysis yielded a fragment with a size of ~ 700 bp.

Plasmid DNA from a single colony was then further restricted with *Bam*HI and *Hind*III to confirm VP1 insertion obtaining a correct size fragment of ~ 650 bp (Fig. 3.3). This clone was subsequently sequenced to confirm VP1 identity (see Annexure for the sequence alignment between the positive control VP1 DNA sequence (pMYE100) and VP1 sequence from transformant). Identities between positive control VP1 DNA sequence from plasmid

pMYE100 and VP1 DNA sequence of transformant was 0.9953052 or 99.5%. Identities between the positive control VP1 amino acid sequence and transformant was 1.0000000 or 100%.

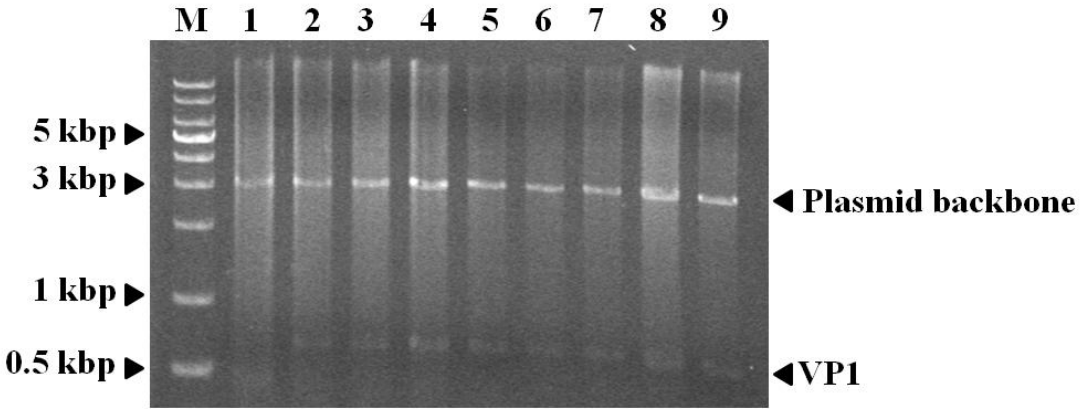


Figure 3.2 *Eco*RI restriction enzyme digest of pGEM®-T-VP1 transformants. M represents the 1 kb DNA ladder. Lanes 1-9 represent isolated and *Eco*RI digested plasmid DNA from the cells of nine different colonies. The VP1 insert size is ~ 700 bp in size.

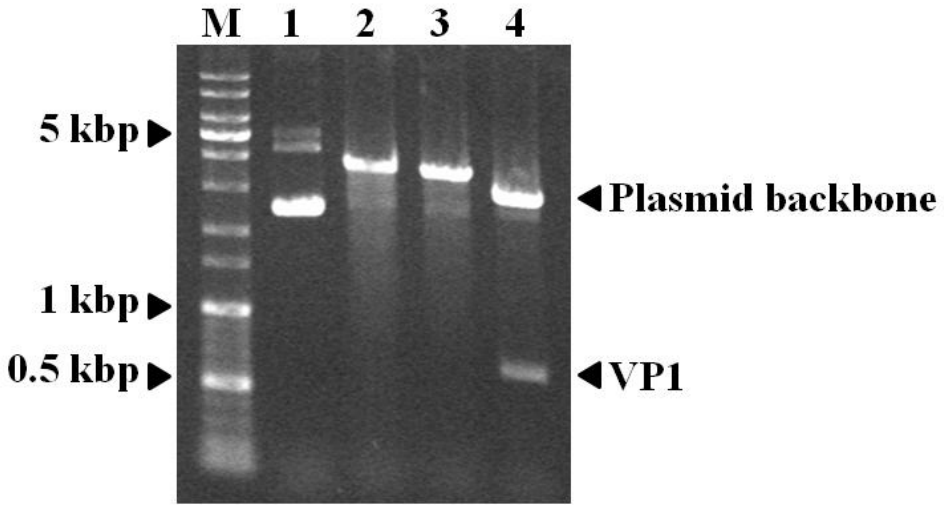


Figure 3.3 *Bam*HI/*Hind*III restriction enzyme digest of a pGEM®-T-VP1 transformant. M represents the 1 kb molecular DNA plus ladder. Lane 1 represents undigested plasmid DNA; lane 2, *Bam*HI digested plasmid DNA; lane 3, *Hind*III digested plasmid DNA and lane 4 represents plasmid DNA digested with *Bam*HI and *Hind*III. The VP1 insert is ~ 650 bp in size.

3.1.2 Cloning VP1 coding sequence into plasmid pQE-30

The VP1 sequence in plasmid pGEM®-T-Easy was restricted with *Bam*HI and *Hind*III and cloned into the bacterial expression vector pQE-30, transferred into *E. coli* (Top 10) cells, and then transferred into *E. coli* M15 cells. VP1 was re-confirmed in M15 transformants using *Bam*HI/*Hind*III restriction enzyme analysis of isolated plasmid DNA yielding a digested fragment with the correct size of ~ 650 bp (Figs. 3.4 and 3.5). Three of the VP1 containing clones were selected for VP1 protein production and cells with plasmid pMYE105 without VP1 insertion was used as a negative control.

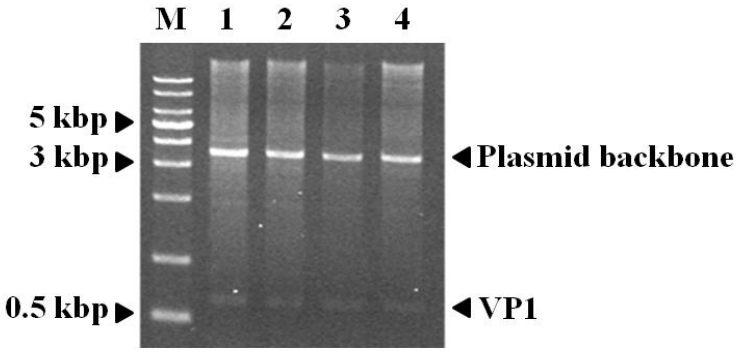


Figure 3.4 *Bam*HI/*Hind*III restriction enzyme digest of plasmid pMYE103 containing VP1 gene sequence derived from *E. coli* (Top 10) cells. M represents the 1 kb DNA ladder; lanes 1-4, plasmid pQE–30 containing the VP1 insert digested with *Bam*HI and *Hind*III. The VP1 insert is ~ 650 bp in size.

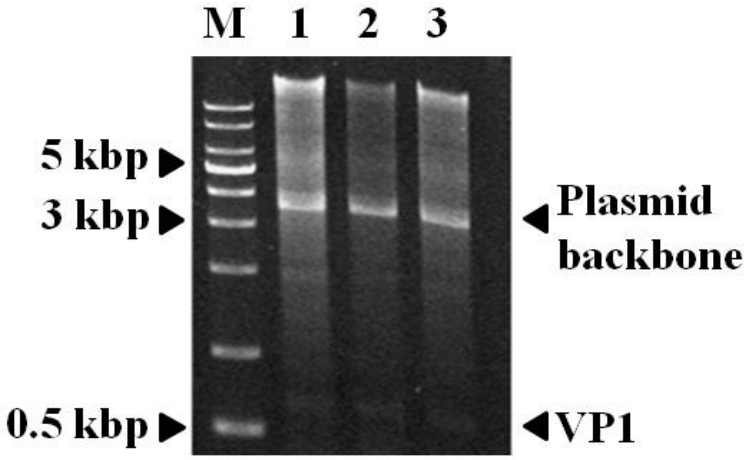


Figure 3.5 *Bam*HI/*Hind*III restriction enzyme digest of plasmid pMYE103 containing VP1 sequence derived from *E. coli* M15 cells. M represents the 1 kb DNA ladder; lanes 1-3, plasmid pQE–30 containing the VP1 insert digested with *Bam*HI and *Hind*III. VP1 insert is ~ 650 bp in size.

3.1.3 Cloning VP1 coding sequence into binary vector pMYV497

VP1 in plasmid pMYE100 was PCR amplified with primers VP1(*Bam*HI)-F and VP1(*Kpn*I)-R yielding a ~ 650 bp amplicon (Fig. 3.6) and cloned into plasmid pGEM®-T-Easy. The presence of VP1 was confirmed by restriction enzyme analysis using *Eco*RI (Fig. 3.7) and digest with enzymes *Bam*HI and *Kpn*I resulted in a ~ 650 bp DNA fragment (Fig. 3.8). A positive clone was then sequenced to confirm VP1 identity (see Annexure for the sequence alignment between the positive control VP1 DNA sequence (pMYE100) and VP1 sequence from transformant).

The VP1 sequence was then cloned into the binary vector pMYV497 and *A. tumefaciens* LBA4404 cells were transformed with the plasmid. Isolated plasmid DNA from putative transformants was used to re-transform *E. coli* (Top 10) to confirm transformation of Agrobacterium cells with VP1 (Figs. 3.9 and 3.10). The plasmid called pMYE108 was subsequently used for transient agro-infiltration of *N. benthamiana* and *N. tabacum* plants.

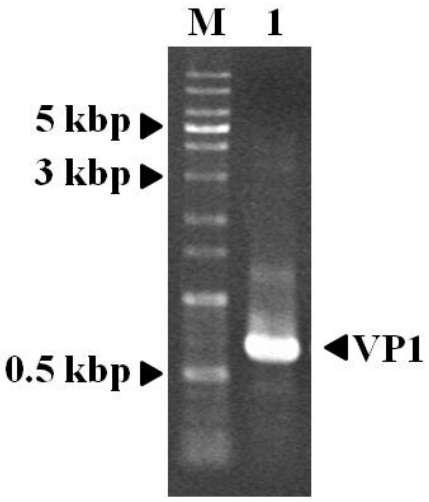


Figure 3.6 Amplification of VP1 coding sequence with primers VP1(*Bam*HI)-F and VP1(*Kpn*I)-R from plasmid pMYE100 template DNA. M represents the 1 kb Molecular DNA plus ladder and lane 1 represents the VP1 amplicon with a size of ~ 650 bp.

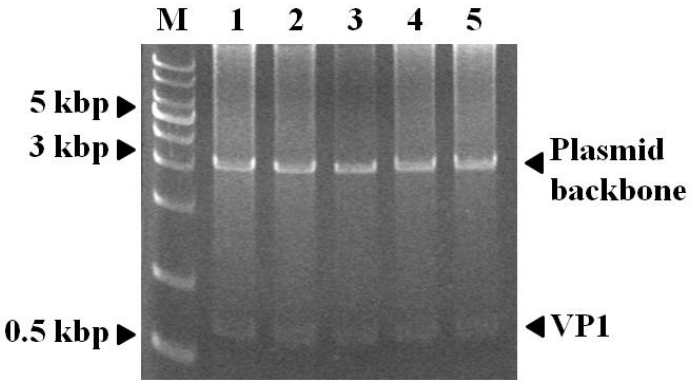


Figure 3.7 *Eco*RI restriction enzyme digest of putative pGEM®T-VP1 transformants derived from *E. coli* (Top 10) cells. M represents the 1 kb DNA ladder marker and lanes 1-5 represent isolated and *Eco*RI digested plasmid DNA from cells of five colonies. The VP1 insert is ~ 650 bp in size.

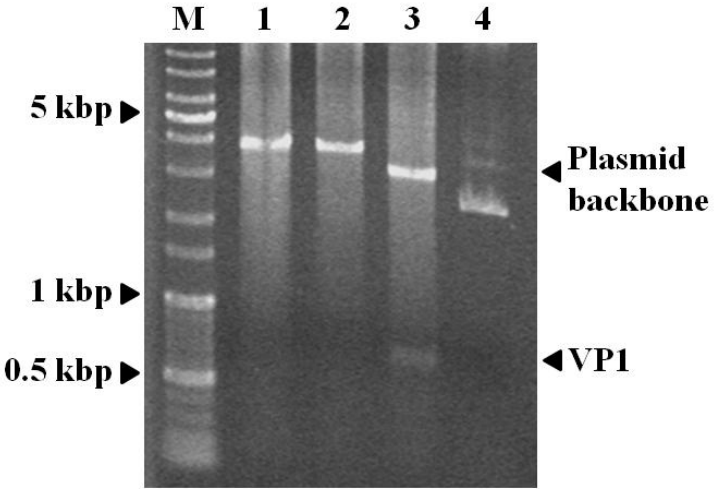


Figure 3.8 *Bam*HI/*Kpn*I restriction enzyme digest of a putative pGEM®T-VP1 transformant derived from *E. coli* (Top 10) cells. M represents the 1 kb molecular DNA plus ladder. Lane 1 represents *Bam*HI digested plasmid DNA; lane 2, *Kpn*I digested plasmid DNA; lane 3, plasmid DNA digested with *Bam*HI and *Kpn*I and lane 4 represents undigested plasmid DNA. The VP1 insert is ~ 650 bp in size.

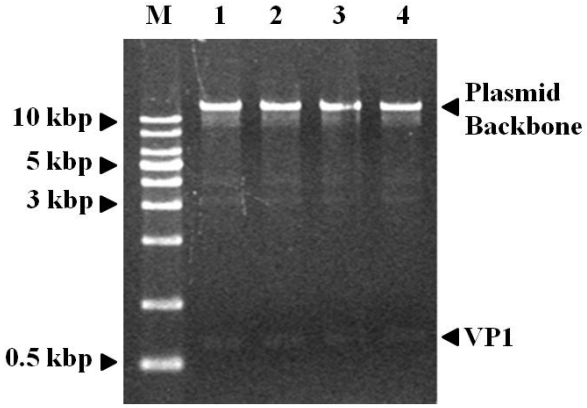


Figure 3.9 *Bam*HI/*Kpn*I restriction enzyme digest of putative pMYV497-VP1 transformants derived from *E. coli* (Top 10) cells. M represents the 1 kb DNA ladder and lanes 1-4 represent plasmid DNA digested with *Bam*HI and *Kpn*I. The VP1 insert is ~ 650 bp in size.

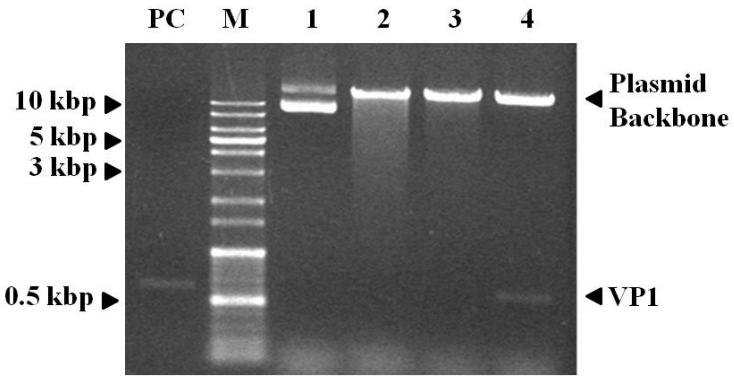


Figure 3.10 *Bam*HI/*Kpn*I restriction enzyme digest of a putative pMYV497-VP1 back-transformant derived from *E. coli* (Top 10) cells. PC represents the positive control of purified VP1 DNA; M, 1 kb molecular DNA plus ladder; lane 1, undigested plasmid DNA; lane 2, *Bam*HI digested plasmid DNA; lane 3, *Kpn*I digested plasmid DNA and lane 4 represents plasmid DNA digested with *Bam*HI and *Kpn*I. The VP1 insert is ~ 650 bp in size.

3.1.4 *In silico* analysis of the VP1 amino acid sequence

Using the NetNGly online glycosylation prediction program, the VP1 protein was predicted to have three N-linked glycosylated sites in its amino acid sequence as indicated below in underlined and red characters on amino acid residues 85, 100 and 103. The CTB signal peptide and SEKDEL ER retention signal are shaded in grey.

CTB signal sequence **VP1 start**

M I K L K F G V F F T V L L S S A Y A H G S T T S T G E S A D P V T A T V E N Y G
G E T Q V Q R R Q H T D V S F I L D R F V K V T P K D Q I N V L D L M Q T P A H T
L V G A L L R T A T Y Y F A D L E V A V K H E G **N** L T W V P N G A P E T A L D **N**
T T **N** P T A Y H K A P L T R L A L P Y T A P H R V L A T V Y N G N C K Y G E G P V
T N V R G D L Q V L A Q K A A R T L P T S F N Y G A I K A T R V T E L L Y R M K R
A E T Y C P R P L L A I H P S E A R H K Q K I V A P V K Q L L G T S E K D E L

VP1 end
ER retention signal

The VP1 protein is susceptible to cleavage by several types of proteases, when the *in silico* peptide cutter tool was applied. Results obtained from the analysis are listed in Table 3.1. *In silico* analysis of the VP1 amino acid sequence revealed that there are two cysteine residues to be exposed to proteolysis by cysteine endopeptidases. NTCB (2-nitro-5-thiocyanobenzoic acid) is a protease that specifically acts on cysteine residues. *In silico* analysis of the VP1 amino acid sequence also revealed that the protein is also susceptible to proteolytic attack from serine proteases.

Table 3.1 Enzymes that cleave the VP1 amino acid sequence (determined using PeptideCutter).

Name of enzyme	No. of cleavages	Positions of cleavage sites
Arg-C proteinase	13	48 49 60 89 136 146 167 179 194 201 204 211 222
Asp-N endopeptidase	9	30 52 58 67 73 96 120 168 240
Asp-N endopeptidase + N-terminal Glu	21	27 30 37 42 52 58 67 73 96 98 104 116 120 159 168 196 205 219 238 240 241
BNPS-Skatole	1	110
CNBr	3	1 76 202
Caspase1	1	97
Chymotrypsin-high specificity (C-term to [FYW], not before P)	19	6 9 10 18 40 56 61 93 94 95 110 129 141 152 158 185 187 200 208
Chymotrypsin-low specificity (C-term to [FYWML], not before P)	53	1 4 6 9 10 13 14 18 20 40 51 56 58 61 73 75 76 81 83 87 88 93 94 95 98 104 108 110 120 129 130 134 137 141 145 148 152 158 170 173 185 187 198 199 200 202 208 213 214 223 234 235 243
Clostripain	13	48 49 60 89 136 146 167 179 194 201 204 211 222
Formic acid	9	31 53 59 68 74 97 121 169 241
Glutamyl endopeptidase	12	28 38 43 99 105 117 160 197 206 220 239 242
Hydroxylamine	2	113 153
Iodosobenzoic acid	1	110
LysC	14	3 5 63 67 103 131 157 176 191 203 224 226 232 240

LysN	14	2 4 62 66 102 130 156 175 190 202 223 225 231 239
NTCB (2-nitro-5-thiocyanobenzoic acid)	2	155 208
Pepsin (pH1.3)	63	3 4 6 8 9 10 12 13 14 17 18 39 40 55 56 57 58 61 72 73 74 75 82 86 87 88 92 93 94 95 97 98 107 108 109 119 120 128 129 137 139 140 151 152 157 158 170 172 173 184 185 186 187 197 198 199 200 207 212 214 233 235 243
Pepsin (pH>2)	45	3 4 6 8 9 10 12 13 14 55 56 57 58 61 72 73 74 75 82 86 87 88 94 95 97 98 107 108 119 120 137 139 170 172 173 184 185 197 198 199 212 214 233 235 243
Proline-endopeptidase [*]	2	212 218
Proteinase K	125	2 4 6 8 9 10 11 12 13 14 17 18 19 23 24 26 30 33 34 35 36 37 40 44 46 52 54 56 57 58 61 62 64 65 70 72 73 75 78 80 82 83 84 86 87 88 90 91 92 93 94 95 96 98 100 101 102 108 109 110 111 115 118 119 120 123 124 127 128 129 132 134 135 137 138 139 141 142 143 147 148 149 150 151 152 158 163 164 166 170 172 173 174 177 178 180 181 183 185 187 189 190 192 193 195 196 198 199 200 205 207 208 213 214 215 216 221 227 228 229 231 234 235 237 243
Staphylococcal peptidase I	12	28 38 43 99 105 117 160 197 206 220 239 242
Thermolysin	72	1 3 5 7 8 9 11 12 13 16 18 29 32 34 36 45 55 56 57 60 61 63 69 71 72 75 79 82 83 85 86 87 90 94 95 100 101 107 118 119 127 133 136 137 146 147 148 150 162 165 171 172 173 176 177 184 188 189 191 194 198 201 204 212 213 214 215 226 227 230 233 234
Trypsin	25	3 5 48 49 60 63 67 89 103 131 136 146 167 176 179 191 194 201 203 204 222 224 226 232 240

3.2 Protein work

3.2.1 Protein expression (bacteria)

3.2.1.1 *Cell growth, protein expression and isolation*

Three clones from plasmid pMYE103 and a single clone of control plasmid pMYE105 were then used for VP1 protein expression analysis in *E. coli* M15. After induction of VP1 expression, protein extracts were separated by SDS-PAGE (Figs. 3.11A and 3.12A). VP1 expression was monitored by immuno-detection analysis with an Anti-His antibody (Fig. 3.11B) and Anti-FMDV antiserum (Fig. 3.12B). A His-VP1 fusion protein with a molecular weight of ~ 25 kDa was detected from all three clones using Anti-His antibody and Anti-FMDV antiserum. Clone #3 was selected for further protein purification using affinity chromatography.

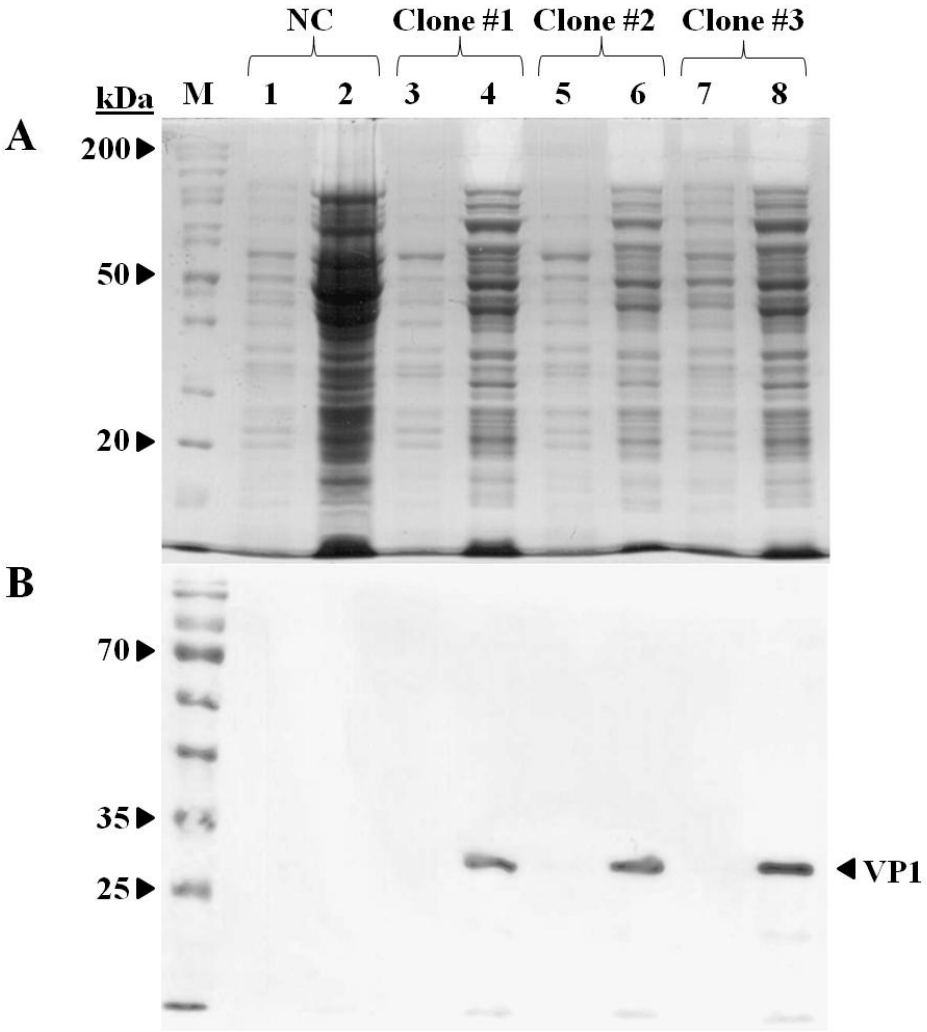


Figure 3.11 Immuno-detection of the Histidine-VP1 fusion protein using an Anti-His antibody after (A) SDS-PAGE analysis and (B) immuno-blotting. M represents the pre-stained protein ladder. NC (lanes 1 and 2) represent un-induced and induced cell fractions from the empty vector negative control plasmid pMYE105. Clones #1, #2 and #3 represent three different clones containing plasmid pMYE103 with VP1 before IPTG induction (lanes 3, 5 and 7) and after IPTG induction (lanes 4, 6 and 8). A His-VP1 fusion protein with a molecular weight of ~ 25 kDa was detected from all three clones using Anti-His antibody.

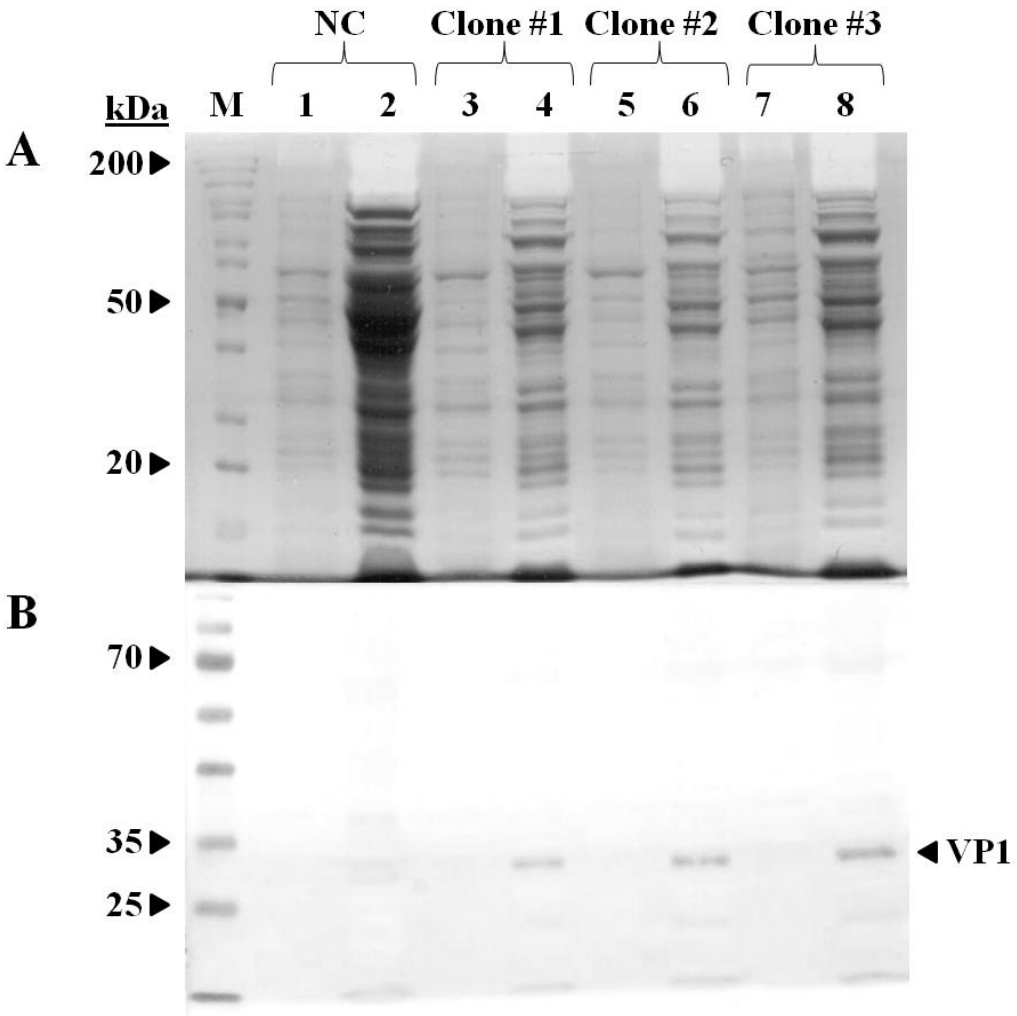


Figure 3.12 Immuno-detection of the Histidine-VP1 fusion protein using Anti-FMDV antiserum after (A) SDS-PAGE analysis and (B) immuno-blotting. M represents the pre-stained protein ladder. NC (lanes 1 and 2) represent un-induced and induced cell fractions from the empty vector negative control plasmid pMYE105. Clones #1, #2 and #3 represent three different clones containing plasmid pMYE103 with the VP1 before IPTG induction (lanes 3, 5 and 7) and after IPTG induction (lanes 4, 6 and 8). A His-VP1 fusion protein with a molecular weight of ~25 kDa was detected from all three clones using Anti-FMDV antiserum.

3.2.1.2 *Protein purification*

In silico analyses (PROTEIN CALCULATOR v3.3) of the VP1 fusion protein revealed the addition of 12 amino acids to the native VP1 peptide sequence increasing the total amount amino acids translated to 225 (from 213) resulting in a ~ 24.92 kDa fusion protein. A His-VP1 fusion protein of ~ 25 kDa was produced in *E. coli* showing immunogenicity against both a His antibody and a polyclonal FMDV antiserum raised against FMDV, Serotype O (Figs. 3.13A and 3.13B). Table 3.2 summarises the yields obtained for three independent elutions. The yield diminishes with each subsequent elution with elution 3 producing the lowest amount. Elution one and two had the highest protein concentrations collectively producing 3.1 mg of purified VP1 fusion protein per 200 mL of bacterial culture or 15.5 milligrams of purified VP1 fusion per litre of bacterial culture. This has been the highest yield obtained in this study for VP1 protein expression. The production and purification of a functional, antigenic VP1 protein within an *E. coli* expression system also then served as a benchmark and positive control in subsequent plant-based transient VP1 agro-infiltration experiments.

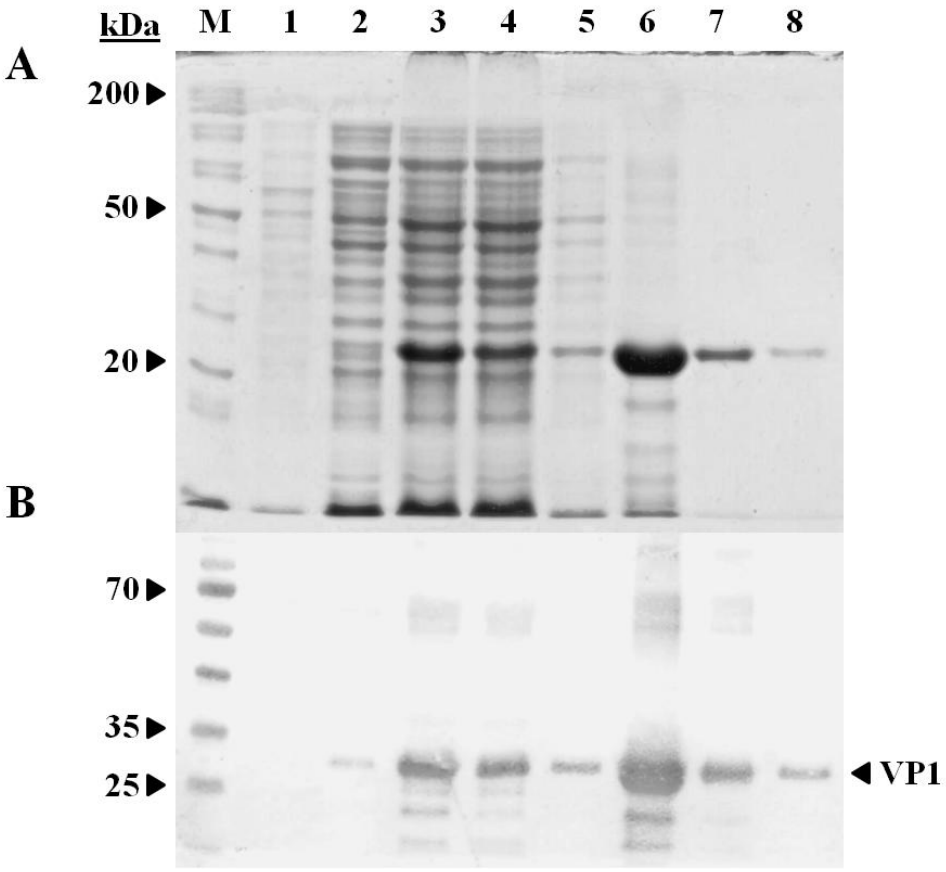


Figure 3.13 Immuno-detection of the Histidine-VP1 fusion protein after purification using Anti-FMDV antiserum. (A) SDS-PAGE analysis and (B) immuno-blotting. M represents a pre-stained protein ladder; lane 1, protein extract from non-induced *E. coli* cells; lane 2, protein extract from IPTG-induced *E. coli* cells; lane 3, proteins in supernatant from lysed bacterial cells; lane 4, flow-through fraction after binding of proteins to affinity column binding His-tagged VP1; lane 5, wash fraction and lanes 6, 7 and 8 represent elution fractions 1, 2 and 3 respectively after eluting the affinity column with imidazole releasing the His-tagged VP1.

Table 3.2 Protein concentrations for the purified His-VP1 fusion protein of Elution fraction's 1 and 2 for three different replicates from the VP1 encoding pMYE103 plasmid.

	R 1 mg/ml	R 2 mg/ml	R 3 mg/ml	Average mg/ml	STD[a]	SE[b]
Elution 1	2.098	2.146	2.815	2.353	0.401	0.231
Elution 2	0.185	0.302	0.672	0.386	0.254	0.147

[a] Standard deviation
[b] Standard error

3.2.2 Protein expression (plants)

3.2.2.1 *Immunodetection of VP1 in infiltrated N. benthamiana leaves*

When *N. benthamiana* leaves were agro-infiltrated with the binary vector pMYE108 for transient expression of VP1, the intensity of the major protein band ~ 50 kDa (Rubisco) greatly decreased over time as shown on an SDS-PAGE gel (Fig. 3.14A). A protein with a size of ~ 30 kDa, very likely representing VP1, could be detected on three, five and seven dpi (days post-infiltration) in protein extracts from agro-infiltrated *N. benthamiana* leaves by immuno-blotting using a polyclonal FMDV antiserum (Fig. 3.14B). On day seven pi, the formation of a smaller sized protein, ~20-25 kDa, could be detected (Fig. 3.14B). This smaller protein was not present in protein fractions of day three and five pi indicating that it was produced on day seven pi and was reactive against the FMDV antiserum. Due to maximal protein expression occurring between day five and seven pi and the possibility of

degradation of the VP1 protein on day seven pi, the same experimental design was applied for subsequent agro-infiltration experiments with *N. tabacum* plants expressing OC-I.

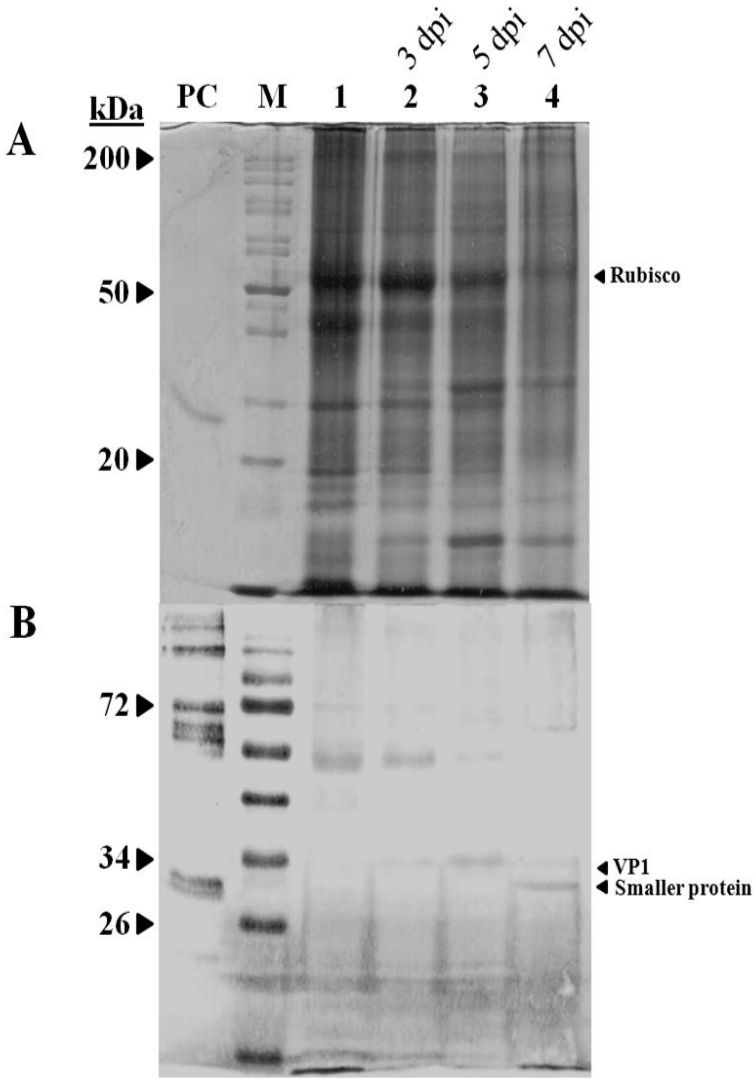


Figure 3.14 SDS-PAGE and immuno-detection of VP1 in *N. benthamiana* leaves agro-infiltrated with the VP1 binary vector pMYE108. (A) SDS-PAGE analysis of leaf protein extracts (~ 50 µg protein) and (B) immuno-detection of the VP1. M represents a "Pageruler" unstained protein ladder; PC, positive VP1 control derived from expression of VP1 in *E. coli* (~ 0.58 µg); lane 1, negative control (un-infiltrated leaf material); lane 2, 3 dpi; lane 3, 5 dpi and lane 4 represents 7 dpi.

3.2.2.2 *Characterization of OC-I transformed N. tabacum leaves*

Plants were first divided into two groups: OC-I expressing and non OC-I expressing controls based on immuno-detection of OC-I, fluorometric GUS activity measurement and cysteine protease activity measurement. A band of ~ 11 kDa representing the correct size for the OC-I protein was detected in *N. tabacum* leaves expressing the OC-I protein using an antiserum raised against OC-I (Fig. 3.15B).

OC-I expressing *N. tabacum* plants were agro-infiltrated with the VP1 binary vector pMYE108. In addition, *N. tabacum* plants were also simultaneously co-infiltrated with the VP1 binary vector pMYE108 and the OC-I binary vector pKYOCI. Table 3.3 summarises total soluble protein concentrations (TSP), activities of cysteine proteases (CP; cathepsin L-like) and GUS activities for both control (non-OC-I expression) and experimental plants (OC-I expressing) agro-infiltrated with *A. tumefaciens* harboring the VP1 plasmid pMYE108. TSP decreased after agro-infiltration in both types of plants. However, control plants (non-OC-I) had further much higher cysteine protease activity than OC-I plants and OC-I plants had in comparison to control plants much higher GUS activity.

The phenotype of infiltrated leaves in control and experimental plants was monitored over the seven day infiltration period (Fig. 3.16) to observe any changes in leaf appearance. In control (non-OCI) leaves, agro-infiltrated with plasmid pMYE108, a more necrotic and chlorotic phenotype was observed on day five and seven pi (Fig. 3.16). These necrotic lesions developed much faster in control plants than in OC-I expressing plants and were more pronounced at day five pi.

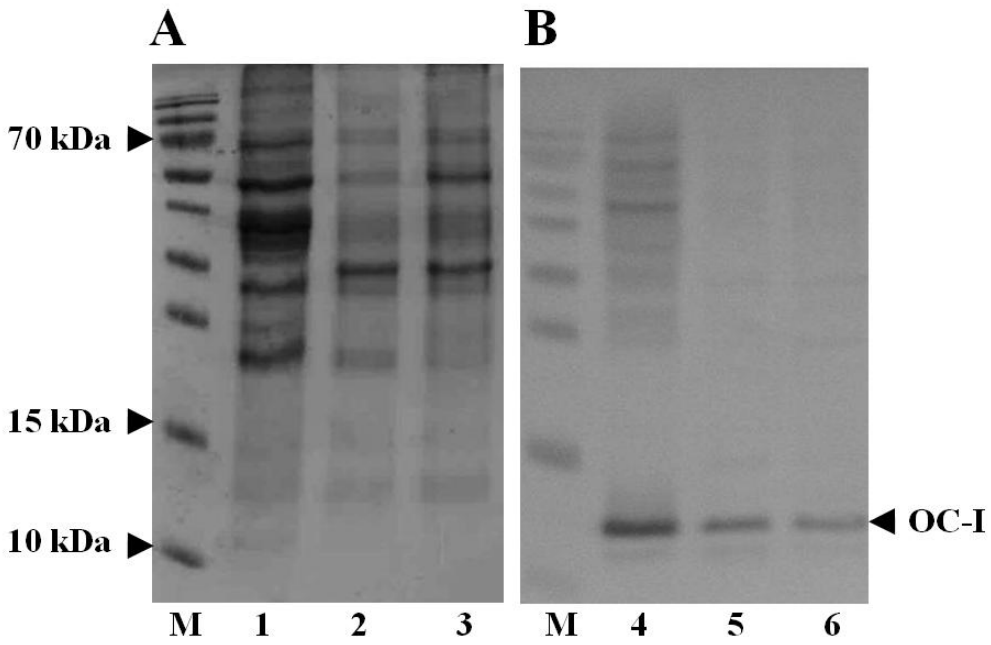


Figure 3.15 SDS-PAGE and immuno-detection of OC-I in OC-I expressing *N. tabacum* leaves agro-infiltrated with the VP1 binary vector pMYE108. (A) SDS-PAGE analysis of tobacco leaf protein extracts (~ 50 µg protein) and (B) immuno-blot analysis for expression of OC-I using an antiserum raised against OC-I. M represents a "Pageruler" prestained protein ladder; lanes 1 and 4, un-infiltrated leaves; lanes 2 and 5, 5 dpi, lanes 3 and 6, 7 dpi.

Table 3.3 Summary of total soluble protein concentrations, activities of cysteine proteases (CP; cathepsin L-like) and GUS in control (non-OC-I expressing) and experimental (OC-I expressing) *N. tabacum* plants agro-infiltrated with the VP1 plasmid pMYE108. Data of control and OC-I expressing plants are the means ±SE of three independent plants.

Plant type	TSP ($mg\ ml^{-1}$)	GUS activity ($ng\ MU\ min^{-1}\ mg\ protein^{-1}$)	CP activity ($FU\ hr^{-1}\ mg\ protein^{-1}$)
Control (non-OC-I expressing)			
UC (Day 1)	8.2 ± 2.1	380.9 ± 125.9	1376.4 ± 143.0*
Day 5 pi	3.7 ± 0.6	653.9 ± 242.4	5630.3 ± 4475.2
Day 7 pi	4.9 ± 0.6	979.9 ± 399.1	9812.2 ± 2955.9
Experimental (OC-I expressing)			
UC (Day 1)	6.3 ± 0.7	3280.7 ± 697.5*	388.9 ± 73.2
Day 5 pi	2.9 ± 0.3	1572.2 ± 275.8*	722.7 ± 419.5
Day 7 pi	4.3 ± 0.4	4345.2 ± 1184.5*	719.7 ± 368.1

UC (Day 1) refers to the un-infiltrated control leaf material measurements conducted on Day one prior to agro-infiltration
* Significant differences between groups were determined via the Student's *t*-test and are indicated by asterisks (t-test; $P < 0.05$)
TSP refers to the total soluble protein content
pi refers to post-infiltration

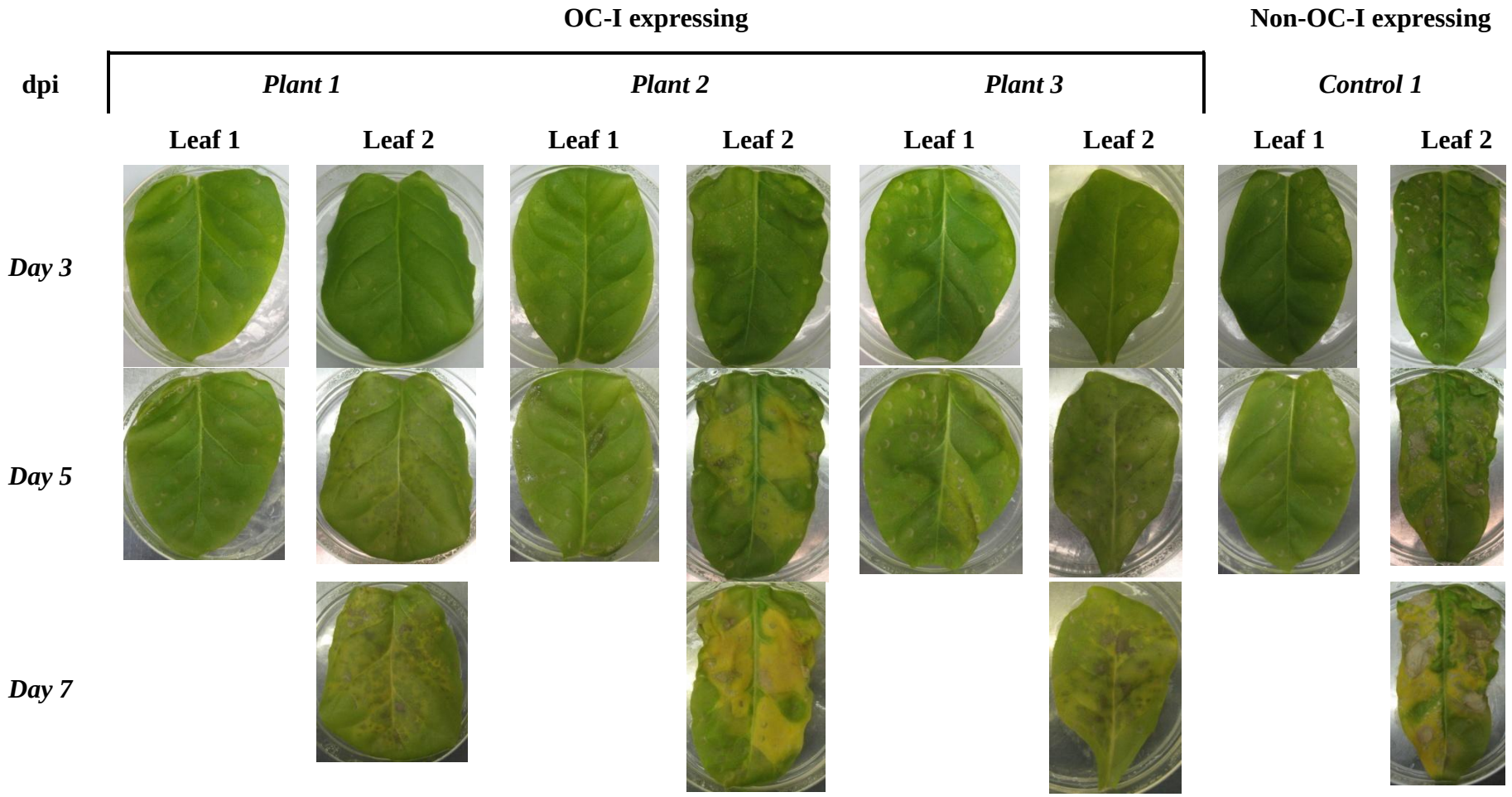
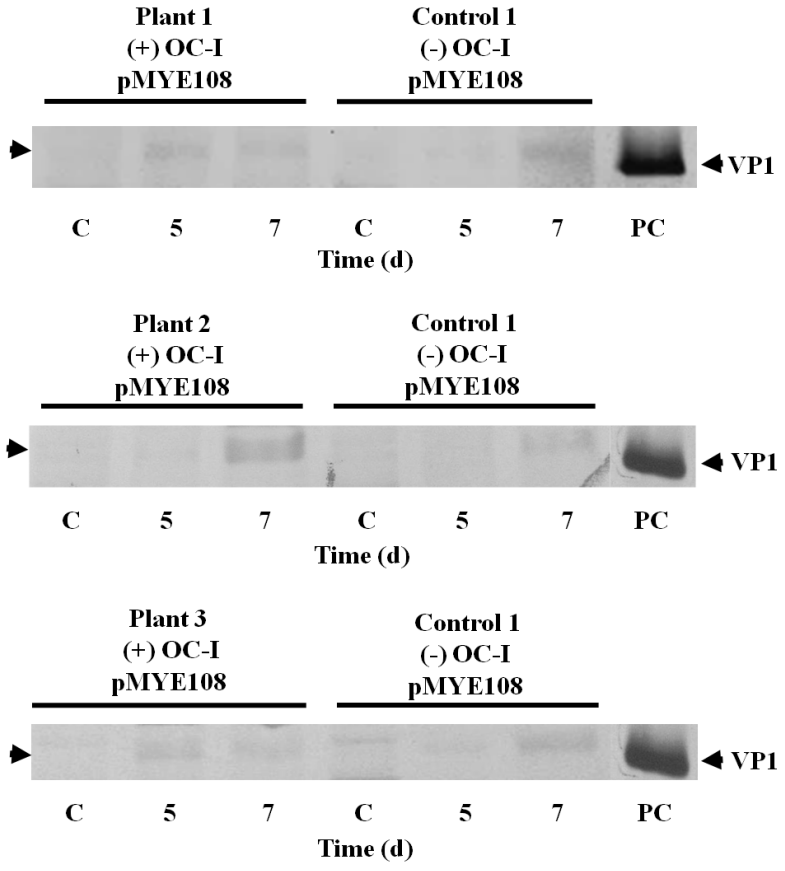


Figure 3.16 Phenotype of OC-I expressing and control *N. tabacum* leaves agro-infiltrated with the binary vector pMYE108. dpi refers to days post infiltration.

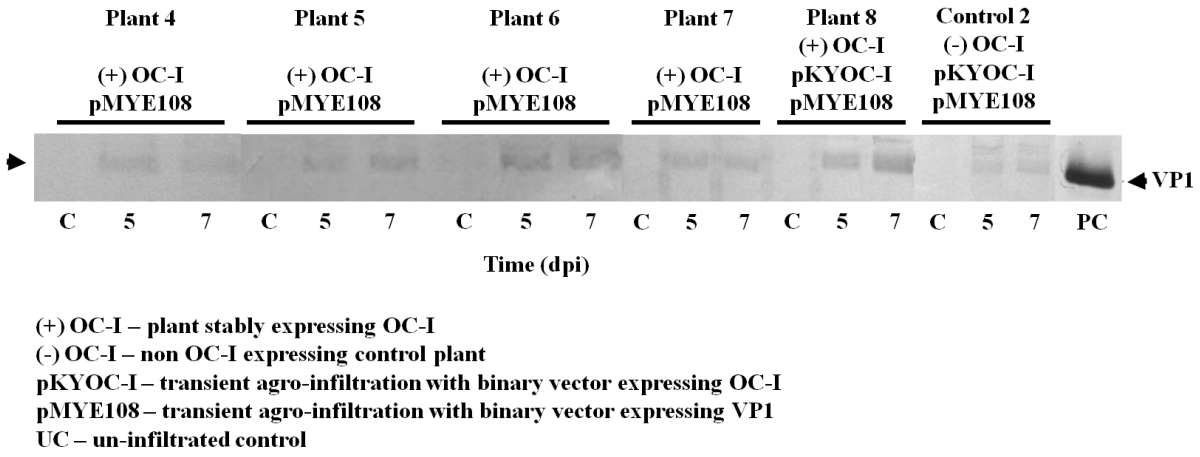
3.2.2.3 *VP1 expression*

N. tabacum plants expressing OC-I and controls were agro-infiltrated with plasmid pMYE108 for transient expression of VP1 and produced an antigenic protein with the size of ~ 30 kDa, larger than the VP1 protein expressed in *E. coli* (positive control; ~ 25 kDa), which was detected on day five and seven pi in all agro-infiltrated leaf tissue (Figs. 3.17 and. 3.18). The protein size was similar to that produced in *N. benthamiana* plants. Experiments were also carried out where OC-I expressing and control *N. tabacum* plants were transiently agro-infiltrated simultaneously with both VP1 and OC-I constructs (pKYOCI and pMYE108). But co-infiltration did not greatly improve VP1 production and was similar to the production also found in one plant (Plant No 6) not co-infiltrated with plasmid pKYOCI (Fig. 3.18).



(+) OC-I – plant stably expressing OC-I
(-) OC-I – non OC-I expressing control plant
pMYE108 – transient agro-infiltration with binary vector expressing VP1

Figure 3.17 Detection by immuno-blotting of VP1 expression in *N. tabacum* plants agro-infiltrated with the VP1 binary vector pMYE108. Transgenic tobacco plants either stably expressed (+) OC-I (Plants 1, 2 and 3) or did not express (-) OC-I (Control 1). Lane PC represents *E. coli*-derived VP1 protein (~ 2.6 µg). Lane C represents un-infiltrated plant material. VP1 expression was measured on day 5 and 7 pi. The position of the *E. coli* expressed VP1 protein of ~ 25 kDa and the position of the putative plant-expressed VP1 (~ 30 kDa) is indicated by arrows.



(+) OC-I – plant stably expressing OC-I
(-) OC-I – non OC-I expressing control plant
pKYOC-I – transient agro-infiltration with binary vector expressing OC-I
pMYE108 – transient agro-infiltration with binary vector expressing VP1
UC – un-infiltrated control

Figure 3.18 Immuno-blotting of VP1 expression in leaves of *N. tabacum* plants agro-infiltrated with only the VP1 binary vector pMYE108 and *N. tabacum* plants co-infiltrated with the VP1 binary vector pMYE108 and OC-I binary vector pKYOCI. Transgenic tobacco plants either stably expressed (+) OC-I (Plants 4 to 8) or did not express (-) OC-I (Control 2). Lanes C represent un-infiltrated leaf material. Lane PC represents *E. coli*-derived VP1 protein (~ 1.3 µg). VP1 expression was measured on day 5 or 7 pi. The position of the *E. coli* expressed VP1 protein is ~ 25 kDa and the position of the putative plant-expressed VP1 (~ 30 kDa) is indicated by arrows.

The cysteine protease activities measured for each plant are shown in Table 3.4. Cysteine protease activities varied within OC-I plants but always had lower activities than the control (non-OC-I). VP1 amounts in agro-infiltrated leaves were determined after immuno-blotting and scanning blots by quantitative densitometry. The highest VP1 amount was detected in *N. tabacum* Plant 2 on day seven pi amounting to 2.95 %TSP but there was no clear relationship between cysteine protease activity measured in various plants and TSP in OC-I expressing plants or between OC-I plants and non-OC-I plants.

Table 3.4 Cysteine proteinases (CP; cathepsin L-like) activities and VP1 content of tobacco plants agro-infiltrated with the binary vector pMYE108 (VP1 construct).

	Plant 1 (+) OCI pMYE108		**Plant 2 (+) OCI pMYE108**		**Plant 3 (+) OCI pMYE108**		**Plant 4 (+) OCI pMYE108**		**Plant 5 (+) OCI pMYE108**		**Plant 6 (+) OCI pMYE108**		**Plant 7 (+) OCI pMYE108**		**Control 1 (-) OCI pMYE108**	
	CP	**%TSP**	**CP**	**%TSP**	**CP**	**%TSP**	**CP**	**%TSP**	**CP**	**%TSP**	**CP**	**%TSP**	**CP**	**%TSP**	**CP**	**%TSP**
UC	13.0	-	559.8	-	246.2	-	198.1	-	723.4	-	0.0	-	513.9	-	1122.3	-
5 dpi	5235.6	0.74	193.2	0.53	2295.7	1.52	0.0	0.44	0.0	0.22	0.0	0.37	0.0	0.53	14580.3	0.58
7 dpi	227.3	0.36	1048.5	2.95	705.4	0.87	0.0	0.27	0.0	0.21	881.9	0.49	0.0	0.43	8370.7	1.00

CP refers to cysteine protease activity measured in FU hr^{-1} mg $protein^{-1}$

UC refers to un-infiltrated leaf material

%TSP refers to the VP1 protein content expressed as a percent of total soluble protein using quantitative densitometry

(+) OCI refers to OCI expressing plants

(-) OCI refers to non-OCI expressing control plants

Plant numbers in table correspond to those in immuno-blots

4. DISCUSSION

In this study, the FMD VP1 protein could be expressed in both a bacterial (*Escherichia coli*) and a plant-based (*Nicotiana benthamiana* and *Nicotiana tabacum*) expression system confirming previous reports of expression of VP1 in these particular expression systems (Kupper *et al.*, 1981; Song *et al.*, 2004a; Andrianova *et al.*, 2011) and VP1 protein expression can exceed 20% of the total cell protein (Neubauer *et al.*, 1992). Proteins of mammalian origin are now being routinely produced in a secreted form with yields in the gram/litre scale (Georgiou and Segatori, 2005). In this regard, 14 mg L^{-1} bacterial culture, as found in this study, is rather low in comparison to other reported fusion proteins. Reasons for this might be the lack of long stretches of similar codons (Makrides, 1996) or consecutive runs of low-usage codons either near the 5'-end (Chen and Inouye, 1990) of the VP1 mRNA or throughout the whole sequence (Bula and Wilcox, 1996). Also the use of plant-codon optimized VP1 might have prevented higher expression as well as His-tagging of VP1. Pérez-Filgueira *et al*. (2004) reported that when VP1 was additionally tagged with the p24 capsid protein of HIV-1, a major component of the gag poly-protein, p24-His yields were 10–15 times higher than those of VP1-His (Pérez-Filgueira *et al.*, 2004). The VP1 amount produced in this study is, however, comparable to the amount previously reported for expression of a β-galactosidase-VP1 fusion protein (Broekhuijsen *et al.*, 1986). Production as a His-tagged fusion protein in *E. coli* has also the advantage of more simple downstream protein enrichment (e.g. immuno-precipitation) and efficient purification.

Unfortunately, bacterial expression systems, although efficient and offering low cost and convenient production (Yin *et al.*, 2007), are limited in their capacity to perform post-translational modifications on complex proteins such as N- and O-linked glycosylation, fatty

acid acylation, phosphorylation or disulfide-bond formation (Yin *et al.*, 2007). These systems are therefore only ideal for expressing unmodified, antigenically active proteins required in high yields. Further, using IPTG, as done in this study, is a major disadvantage for the induction of protein expression. Application of IPTG in a large scale production greatly increases production cost. Therefore, there is a need for alternative lower cost inducers, such as lactose, to replace IPTG in a bacterial system (Neubauer *et al.*, 1992). With regards to VP1 expression, yields obtained in *E. coli* in this study would also be the benchmark to exceed. Future work has therefore also to entail a more rigorous purification of the VP1 fractions and conducting X-ray crystallography to determine possible changes in VP1 structure (Svergun *et al.*, 2001) and testing antigenicity of the purified VP1 protein in an animal model (Griffin, 2002).

Expressed VP1 was found in the soluble fraction possibly due to retention of VP1 in the ER where many of the resident luminal proteins are soluble (Munro and Pelham, 1987). However, expression of VP1 in a plant-based system was by far less efficient when compared to VP1 expression in the *E. coli* system. This was despite expressing VP1 in the ER which can greatly minimize proteolytic degradation (Conrad and Fiedler, 1998; Wydro *et al.*, 2006; Benchabane *et al.*, 2008). Using immuno-blotting for detection of a correct size VP1 protein, expressed either in *N. benthamiana* (0.31 %TSP on day five pi, data not shown) or *N. tabacum* (on average 0.75 %TSP), was further affected by the antiserum used exhibiting non-specific binding to various plant proteins. This limited the overall accuracy of VP1 quantification. Using immuno-detection for determination of the amount of protein expressed, as done in this study, can therefore only be regarded as a rather crude determination method very much dependent on the quality of the antiserum.

When the VP1 protein was expressed in either of the two tobacco types, a higher molecular weight size protein than expected for VP1 was detected after immuno-blotting. Increase in VP1 size might have been caused by a change in the protein's tertiary structure. Three N-glycosylated sites in the VP1 amino acid sequence amino acid residues were predicted at positions 85, 100 and 103 when the NetNGly tool kit was applied. *In silico* analyses revealed that the B cell epitopes of the VP1 protein, which are amino acids 21-40 (Song *et al.*, 2004b), 135-160 (Zamorano *et al.*, 1994) and 200-213 (van Lierop *et al.*, 1992), are not shielded by sugar residues allowing N-linked glycosylation and detection (Fritschy, 2008; Mikschofsky *et al.*, 2009). Therefore, targeting VP1 to the secretory pathway and retention to the ER sub-cellular compartment in *N. benthamiana* may have pre-disposed VP1 to post-translational modifications, such as N-glycosylation (Gomord *et al.*, 2010), attaching sugar chains to these glycosylation sites (Ueda *et al.*, 2000). This might have resulted in a molecular weight shift of VP1 from 25 kDa (Kurz *et al.*, 1981) to the detected 30 kDa in this study. Increases in molecular weight were also previously found in transgenic potato plants expressing the CTB protein (Arakawa *et al.*, 1997) as well as in tobacco seeds expressing phaseolin (Pueyo *et al.*, 1995). Sriraman *et al.* (2004) also reported a slight increase in the molecular weight of the heavy and light chains of the mouse/human chimeric IgG1 antibody when expressed in *N. tabacum* plants. This is due to a lower electrophoretic mobility in SDS-PAGE. An increase in molecular weight was also detected for BVDV (bovine viral diarrhoea virus) glycoprotein E2 when expressed in alfalfa and the regulatory sequences, TEV-L and SEKDEL, were still attached (Dus Santos and Wigdorovitz, 2005). The failure of the tobacco plant cell to remove such leader peptide might therefore have resulted in an increased protein molecular weight. Extended characterisation is therefore still required for the system used. This should, for example, include PNGase F treatment and MALDI-TOF-MS analysis confirming N-linked glycosylation (Triguero *et al.*, 2011).

In agro-infiltrated *N. benthamiana* plants, the possible degradation of VP1 (or alternatively the production of another protein), with the appearance of a smaller molecular weight protein, occurred on day seven pi. Appearance of a smaller size protein has also been found in *E. coli* when expressing a β-galactosidase-VP1 fusion protein (Broekhuijsen *et al.*, 1986). This might suggest that sequences are truncated by protease action over time. Limited proteolysis in *N. sylvestris* plants has been reported when expressing class I chitinases during later stages of post agro-infiltration (6-10 dpi) (Schöb *et al.*, 1997). A similar process might have also occurred in tobacco plants resulting in lower VP1 yield due to protease action inside the cell or after cell disruption during harvest. *In silico* analysis of the amino acid sequence of VP1 revealed that there are two cysteine residues which are exposed to proteolysis by cysteine endo-peptidases. The VP1 protein's CTB signal peptide might have been properly processed, but the VP1 protein could have been cleaved at amino acid position 155 by a plant-derived cysteine protease and still underwent N-linked glycosylation. The theoretical protein size produced would be ~23 kDa. This was about the same size as the smaller protein produced on day seven pi in *N. benthamiana*. Also, restoration of the native protein due to loss of any N-glycosylated moieties that may have been added to the VP1 cannot be ruled out. This would result in a protein with a faster electrophoretic mobility on a SDS-PAGE gel yielding the determined VP1 protein size (Sriraman *et al.*, 2004).

However, OC-I expressing plants consistently produced the VP1 protein without appearance of a smaller size protein band on day seven pi. The sub-cellular compartment where OC-I is located is the cytosol (Prins *et al.*, 2008). It has yet to be investigated if OC-I location in the cytosol might also protect, to a certain degree, degradation of VP1 produced in the ER. The inhibitor might have also acted on cysteine proteases after releasing VP1 during harvest preventing direct action of proteases on released VP1. Further, it is still unclear if OC-I

expression only affects VP1 degradation due to cysteine protease action. There is also evidence that OC-I expression causes generally a higher total protein content in plants but the exact reason is not yet known (Van der Vyver *et al.*, 2003; Demirevska *et al.*, 2010). However, except for one transformed plant, no higher VP1 production was found in agro-infiltrated OC-I plants and there was also no direct relation between cysteine protease activity and VP1 production. So far, results are therefore not conclusive regarding the exact protective role of OC-I for improving VP1 stability and increased production. A more detailed comparative study is therefore required with a much greater number of plants to demonstrate if constitutive expression of OC-I in transformed plants or transient co-expression of OC-I indeed impacts effectively VP1 expression and stability and where a possible protection occurs.

VP1 expression in the plant system might also have been affected by post-transcriptional gene silencing (PTGS) an integral feature of Agrobacterium-based transient gene expression (Johansen and Carrington, 2001; Voinnet *et al.*, 2003). This greatly reduces the efficiency of gene expression. There are two optimal post-infiltration harvest time-points when carrying out transient agro-infiltration which might not be affected by PTGS (Voinnet *et al.*, 2003; Cazzonelli and Velten, 2006). These are 60 - 72 hr and 5-6 days for proteins post agro-infiltration. The 5-6 days harvest point has been found to be effective for the Green fluorescent protein (Sainsbury *et al.*, 2009) as well as for therapeutic antibodies, such as TheraCIM®, which can be still detected after seven days post agro-infiltration (Rodríguez *et al.*, 2005). However, it has still to be confirmed for this study that tobacco indeed follows the optimal time-point expression at 5-6 dpi.

The agro-infiltration process further activates *PR* (pathogen-related) genes (Yang *et al.*, 2000). Cellular damage occurs at the site of syringe contact (Cazzonelli and Velten, 2006) which could be partly prevented by expression of OC-I. This reaction is known to be problematic for Agrobacterium-mediated transient assays of Solanaceous species (Wroblewski *et al.*, 2005). The yellowing of leaf material and chlorotic/necrotic phenotype is therefore most likely related to the Agrobacterium-infiltration and amount of bacterial cells used. Yellowing possibly arises from bacterial proteins that are transferred to the plant cell by the Type IV secretion system. When using Agrobacterium strain LBA4404, an optical cell density at 1.0, or above, results in leaf yellowing (Wroblewski *et al.*, 2005) and the areas of strongest infiltration are the spongy mesophyll, palisade mesophyll, guard cells, and stomata (Cazzonelli and Velten, 2006). A defence response by tobacco plants occurs, which is most likely a type of hypersensitive response (HR), where the tissue surrounding the infection site becomes necrotic (Hammond-Kosack and Jones, 1996). Increases in VP1 content between day five and seven pi were observed associated with chlorosis as opposed to necrosis. This suggests that chlorosis is still more favourable for protein production as opposed to necrosis where proteins are strongly broken down (Solomon *et al.*, 1999). Transient agro-infiltration of transgenic tobacco, stably expressing the resistance gene *Cf-9*, with a matching avirulence gene *Avr9* caused severe necrosis where leaf tissue collapsed at one dpi and developed into a yellow-brown sector by seven dpi (Hoorn *et al.*, 2000). This might also suggest that the non-native nature of transiently expressed VP1 protein might be toxic to the plant inducing the observed chlorotic/necrotic response.

In summary, this study has shown that VP1 can be produced using either a bacterial or plant-based production system. Currently, the bacterial system seems to be superior to the plant-based system regarding the amount of VP1 produced. With regards to plant-based production

of VP1, the use of transgenic OC-I expressing *N. tabacum* plants might offer benefits regarding stability of expressed VP1. Since VP1 expression and stability is influenced by protease activity and ER retention of the VP1 target protein, the role of other proteases in VP1 expression and direction of protease inhibitors to different cellular compartments might be also studied in the future in much greater detail. For example, in tobacco leaves there are at least three different classes of proteases (cysteine, serine, and aspartic proteases). An *in silico* study has already shown that there are five serine residues that will be exposed to attack from serine proteases. It might therefore be worthwhile to also co-infiltrate the plant host with a serine protease inhibitor, to block serine proteases, in combination with a cysteine protease inhibitor in different cellular localization in both the cytosol and the ER sub-cellular compartment. Also an interesting aspect to investigate might be the establishment of a more efficient transient expression platform based on the Cowpea mosaic virus (CPMV) as well as expressing the entire P1 region from FMDV, encoding all viral capsid proteins, in order to generate virus-like particles (VLPs) with improved stability and immunogenicity. Future research on improving VP1 expression in a plant based system might finally also include co-infiltration of a silencing suppressor, such as p19 (Voinnet *et al.*, 2003) or P1/HC-Pro (Johansen and Carrington, 2001) to improve VP1 transient expression, or the application of a non-tobacco based system for VP1 production.

ANNEXURE

I. PLASMID MAPS

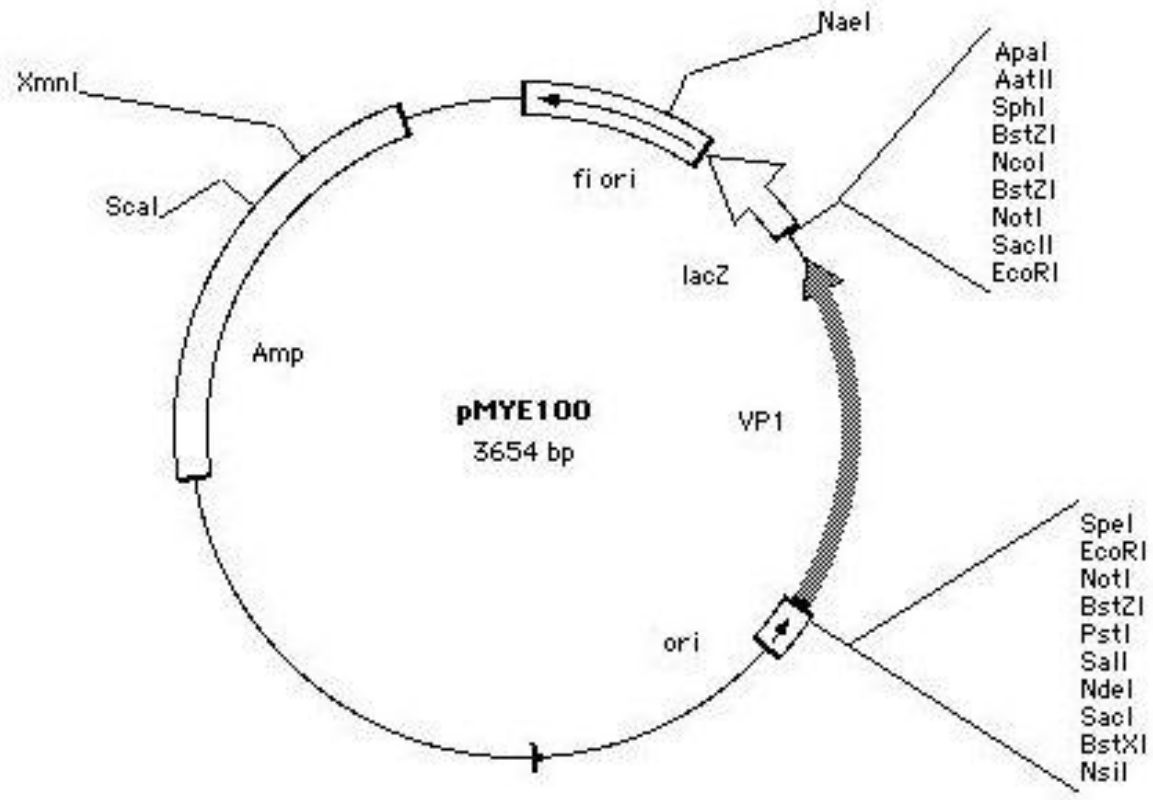


Figure 5.1 Plasmid vector map for pMYE100 containing the VP1 gene in pGEM®T – Easy vector.

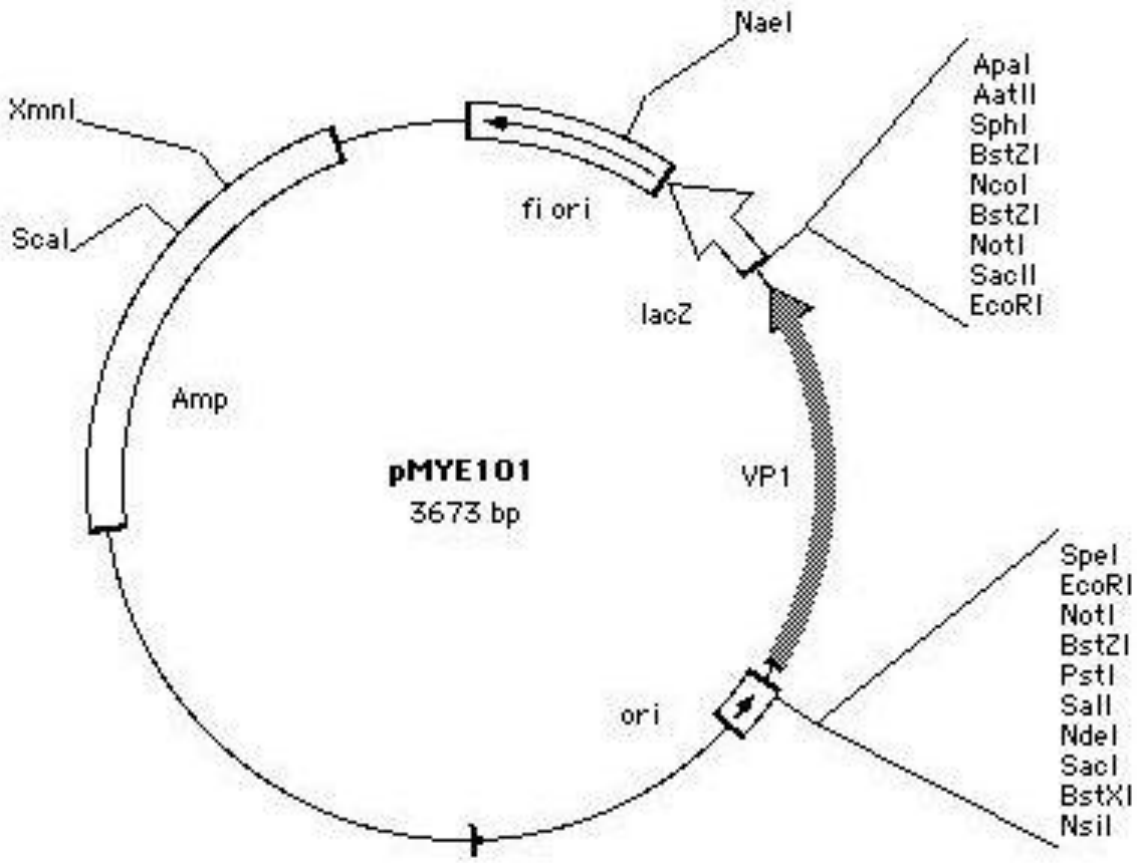


Figure 5.2 Plasmid vector map for pMYE101 containing the VP1 gene with *Bam*HI and *Hind*III restriction enzyme sites and stop codon TAA in pGEM®T – Easy Vector.

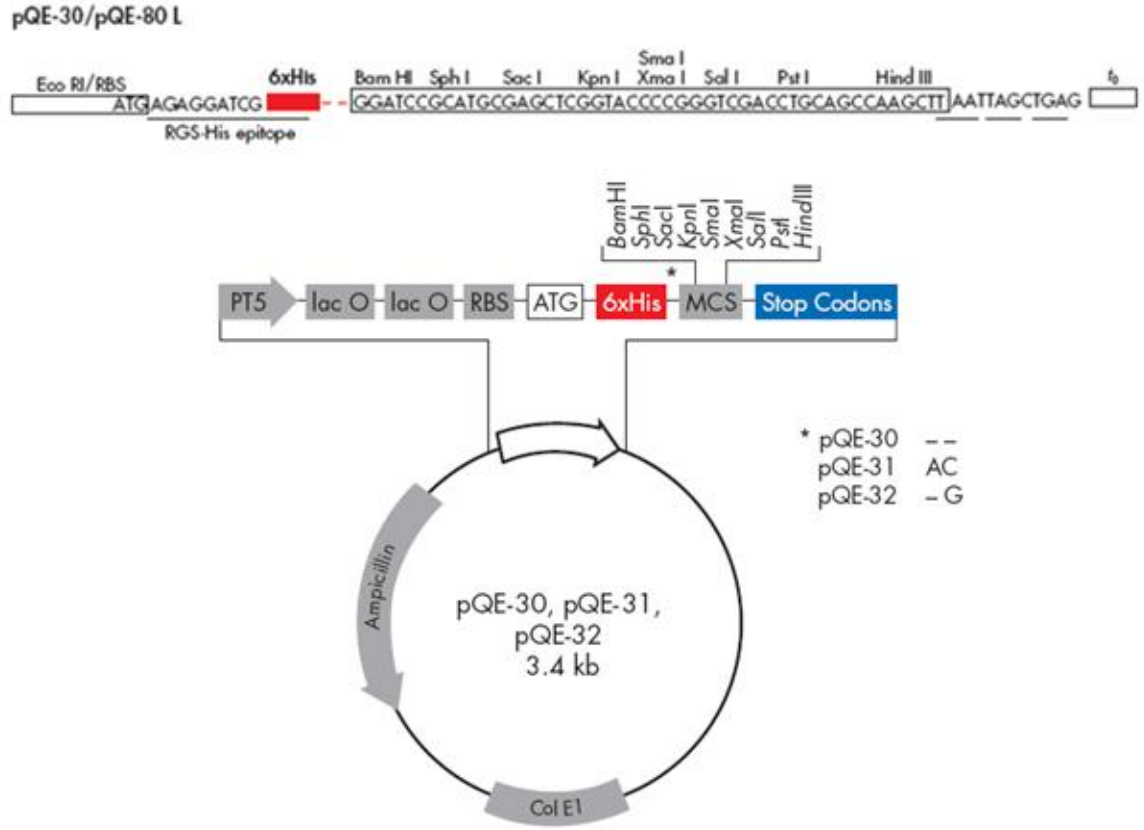


Figure 5.3 Plasmid vector map for pQE-30 bacterial expression vector (adapted from QIAexpressionist™).

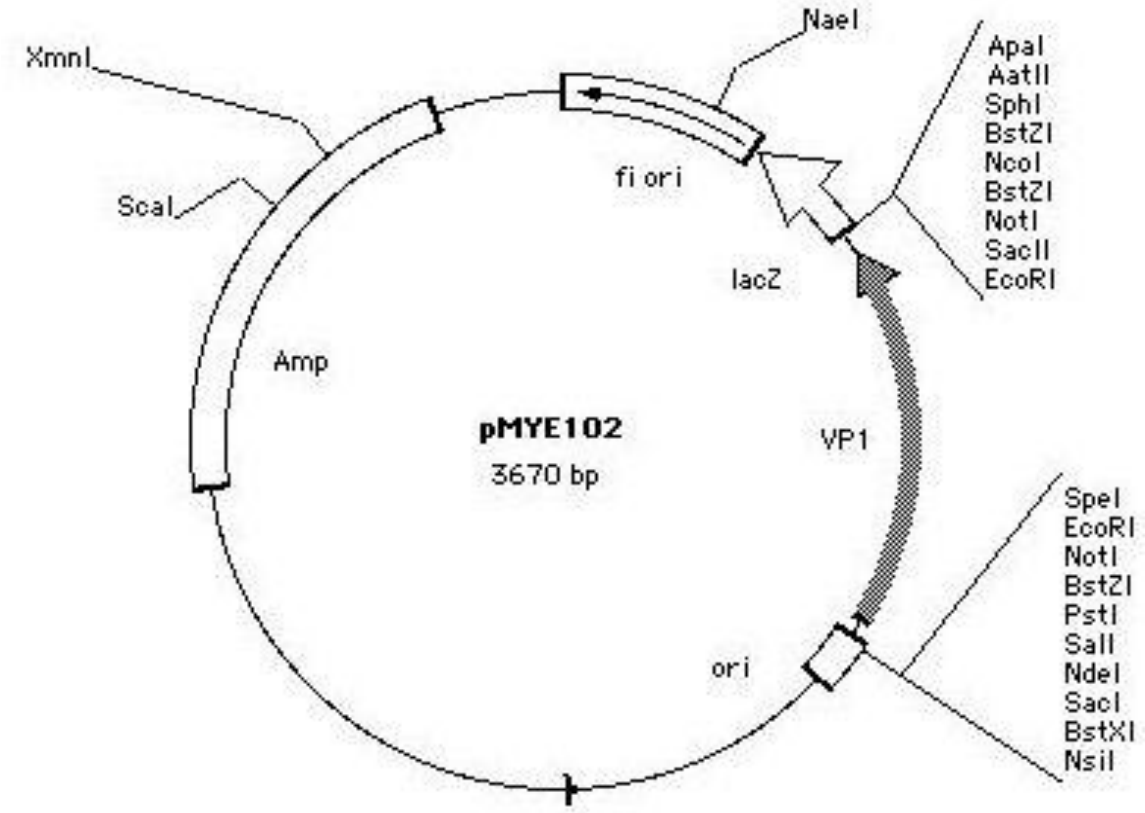


Figure 5.4 Plasmid vector map for pMYE102 containing the VP1 gene with *Bam*HI and *Kpn*I restriction enzyme sites in pGEM®T – Easy Vector.

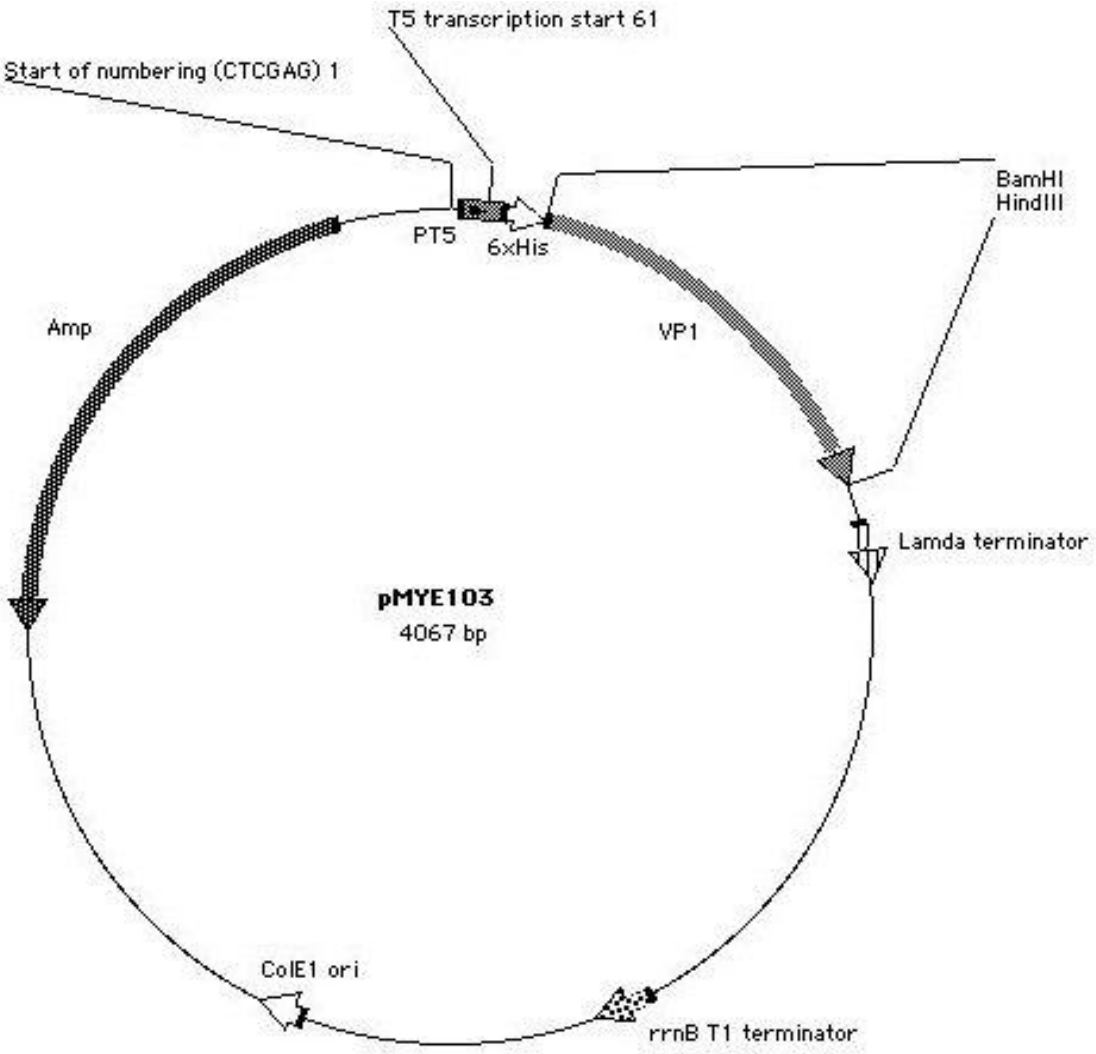


Figure 5.5 Plasmid vector map for pMYE103 containing the VP1 gene with *Bam*HI and *Hind*III restriction enzyme sites and stop codon TAA in bacterial expression vector pQE-30.

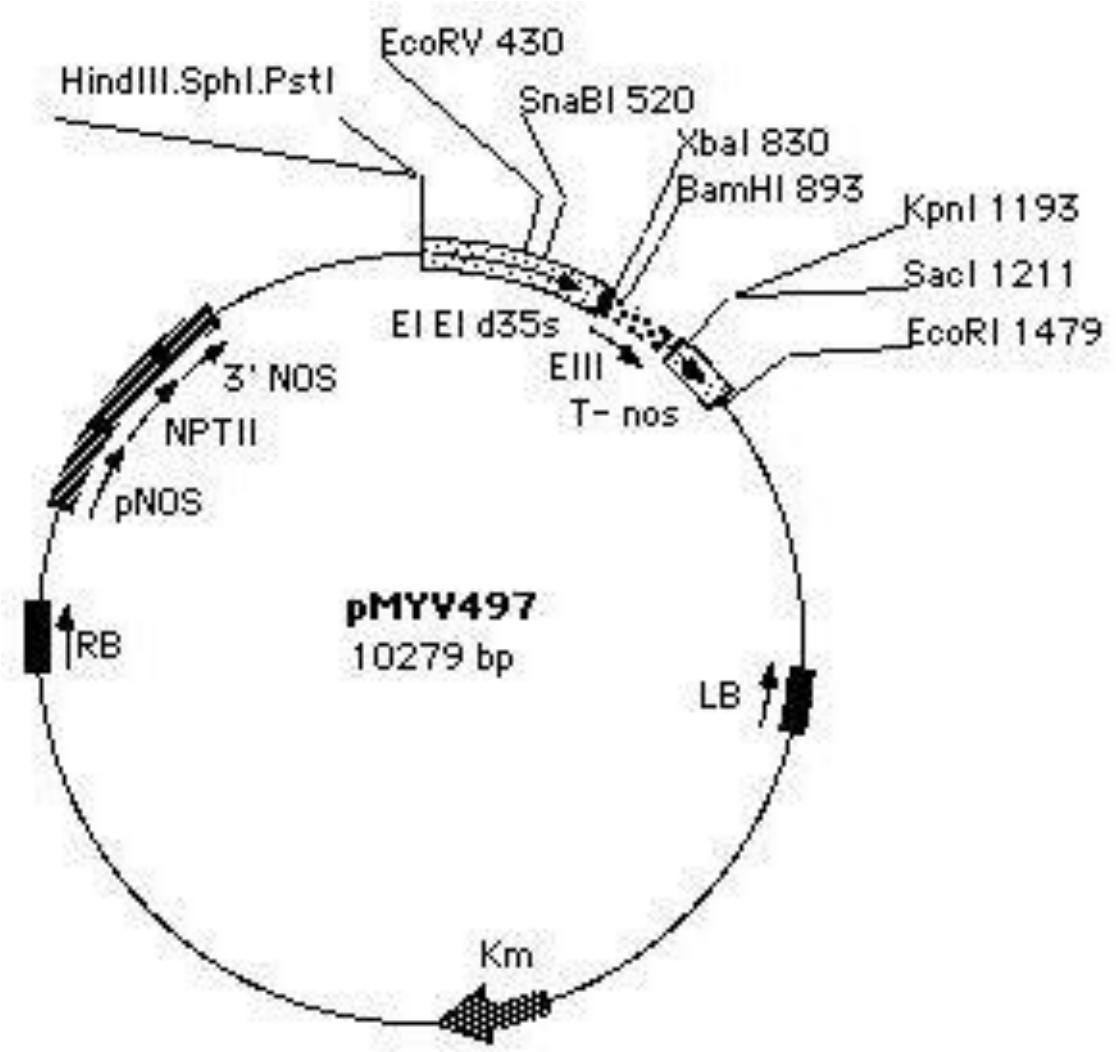


Figure 5.6 Plasmid vector map for pMYV497 (binary vector) containing the CTB signal peptide, Dengue E gene cloned within *Bam*HI and *Kpn*I restriction enzyme sites and ER retention signal SEKDEL.

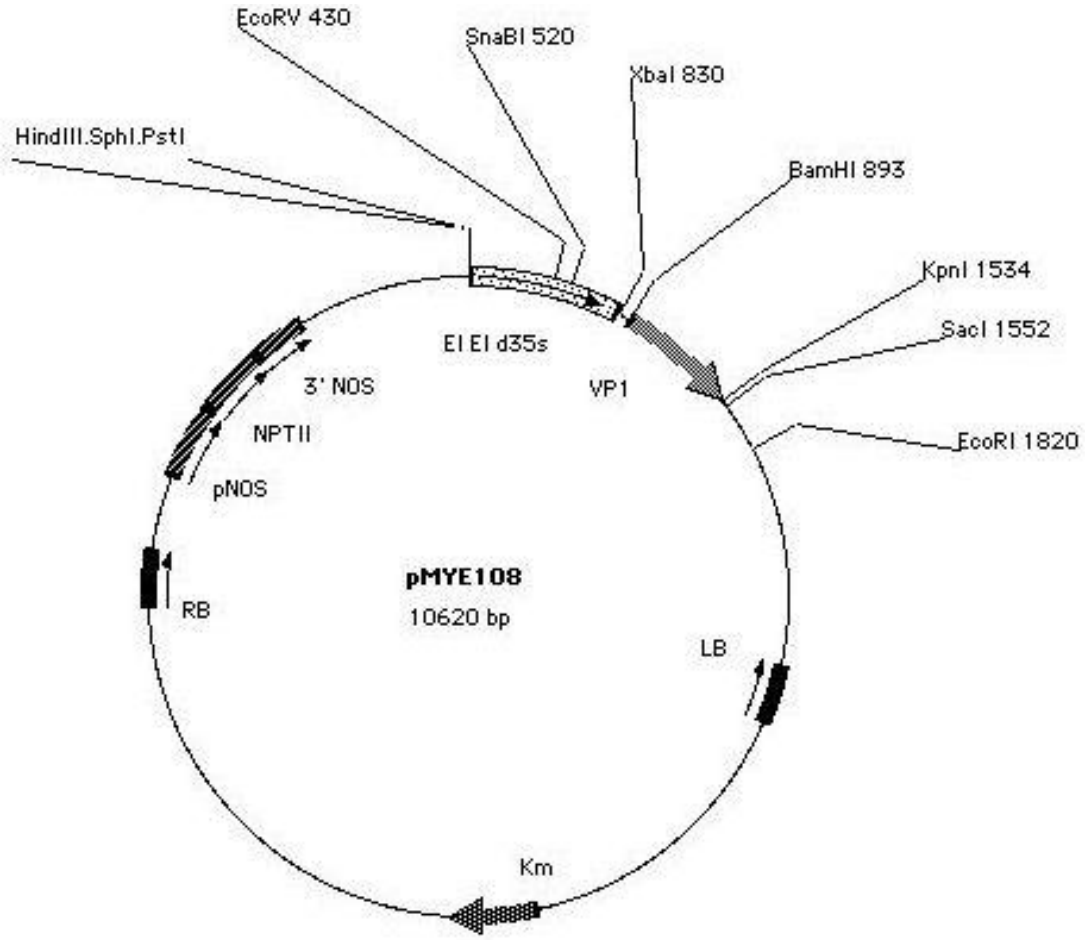


Figure 5.7 Plasmid vector map for pMYE108 containing the VP1 gene within the *Bam*HI and *Kpn*I restriction enzyme sites of the binary vector pMYV497.

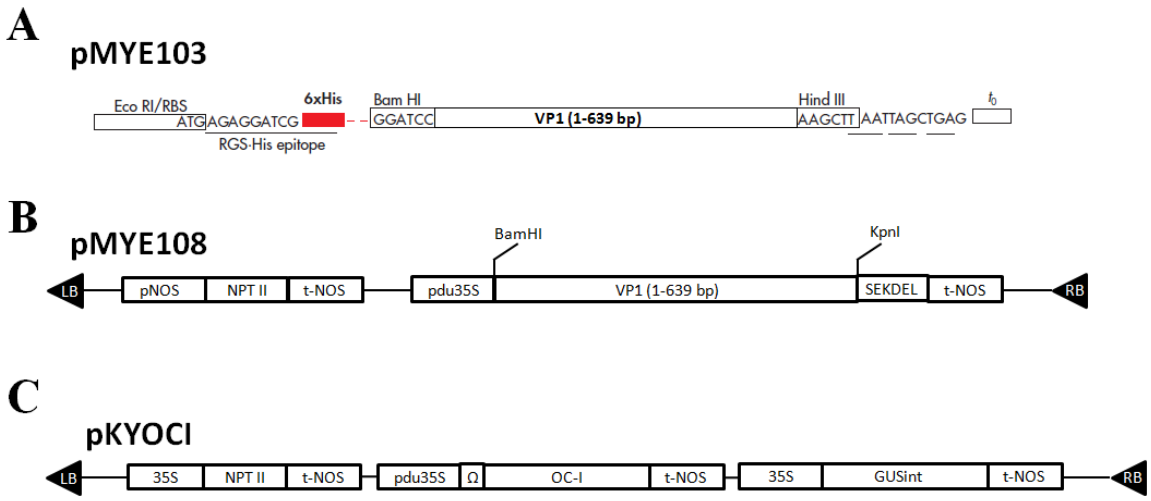


Figure 5.8 Schematic representation of the pMYE103 construct and T-DNAs from binary vectors pMYE108 and pKYOCI. (A) pMYE103 cloned into pQE-30 (Qiagen, Valencia, CA) backbone fused with a 6xHis tag. (B) pMYE108 contains the VP1 gene fused with an ER retention sequence (SEKDEL) under the control of the duplicated Cauliflower Mosaic Virus 35S promoter (pdu35S) and an NPT II (neomycin phosphotransferase II) expression cassette for kanamycin selection of transgenic plants. The pNOS promoter and tNOS terminator are from the *A. tumefaciens* nopaline synthase gene. (C) pKYOCI encodes OC-I coding sequence under the control of a double 35S promoter (P70) from Cauliflower mosaic virus (CaMV) between the left border (LB) and right border (RB). Present on the T-DNA is an Ω leader sequence for gene expression enhancement. The construct also contains the NPT II gene under the control of a 35S promoter (P35SNPTII) used as a selectable marker, and an intron-containing gus gene (GUS) encoding β-glucuronidase (P35S GUSint) under the control of a 35S promoter. Single lines represent un-translated regions.

II. SEQUENCE ALIGNMENTS

a) Alignment : VP1 gene sequence (Genbank) and codon-optimized VP1 gene sequence from pMYE100

```
                    10        20        30        40        50        60
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native      1  ACCACCTCCACAGGTGAGTCGGCTGACCCCGTGACTGCCACTGTTGAGAACTACGGTGGT
Optimal     1  ACCACCTCCACAGGTGAGTCCGCTGATCCTGTGACTGCCACTGTGGAGAACTACGGTGGT

                    70        80        90       100       110       120
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     61  GAGACACAGGTCCAGAGACGCCAACACACGGATGTCTCGTTCATATTAGACAGATTTGTG
Optimal    61  GAGACACAGGTCCAGAGACGCCAGCACACCGATGTCTCCTTCATCTTGGACAGGTTCGTG

                   130       140       150       160       170       180
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     121 AAAGTAACACCAAAAGACCAAATTAATGTGTTGGACCTGATGCAAACCCCTGCACACACT
Optimal    121 AAGGTGACACCGAAGGACCAGATCAACGTGTTGGACCTGATGCAGACCCCTGCACACACC

                   190       200       210       220       230       240
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     181 TTGGTAGGCGCGCTCCTCCGTACTGCCACCTACTACTTCGCAGATCTAGAAGTGGCAGTG
Optimal    181 TTGGTCGGAGCTCTCCTCCGCACTGCCACCTACTACTTCGCAGATCTCGAGGTGGCAGTG

                   250       260       270       280       290       300
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     241 AAACACGAGGGGAACCTTACCTGGGTCCCGAACGGGGCGCCCGAGACAGCGTTGGACAAC
Optimal    241 AAGCACGAAGGCAACCTCACCTGGGTCCCTAACGGAGCTCCAGAGACAGCCTTGGACAAC

                   310       320       330       340       350       360
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     301 ACCACCAATCCAACGGCCTACCACAAGGCACCGCTCACCCGGCTTGCACTGCCTTACACG
Optimal    301 ACCACCAATCCAACCGCCTACCACAAGGCACCTCTCACCCGGCTCGCACTGCCTTACACC

                   370       380       390       400       410       420
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     361 GCACCACACCGTGTCTTGGCTACTGTTTACAACGGGAACTGCAAGTATGGCGAGGGCCCC
Optimal    361 GCACCACACCGCGTCTTGGCCACCGTGTACAACGGCAACTGCAAGTACGGCGAGGGCCCA

                   430       440       450       460       470       480
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     421 GTGACCAATGTGAGAGGTGACCTGCAAGTATTGGCCCAGAAGGCGGCAAGAACGCTGCCT
Optimal    421 GTGACCAACGTGAGAGGTGACCTGCAGGTGCTCGCCCAGAAGGCCGCCAGGACACTGCCT

                   490       500       510       520       530       540
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     481 ACCTCCTTCAACTACGGTGCCATCAAAGCCACTCGGGTGACTGAACTGCTTTACCGCATG
Optimal    481 ACCTCCTTCAACTACGGAGCCATCAAGGCCACTCGGGTGACTGAGCTCCTCTACCGCATG

                   550       560       570       580       590       600
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     541 AAGAGGGCCGAAACATACTGTCCCCGGCCTCTTTTGGCTATTCACCCGAGCGAAGCTAGA
Optimal    541 AAGAGGGCCGAGACCTACTGCCCACGGCCTCTCTTGGCCATCCACCCAAGCGAGGCCAGA
```

```
                        610        620        630
                ....|....|....|....|....|....|....|....
Native      601 CACAAACAAAAGATTGTGGCGCCTGTGAAACAGCTTCTG
Optimal     601 CACAAGCAGAAGATCGTGGCTCCTGTGAAGCAGCTCCTC
```

Native (top gene sequence) refers to the VP1 gene sequence from Genbank (Accession number - AF428246). Optimal (bottom gene sequence) refers to the codon-optimized VP1 coding sequence within plasmid pMYE100. Cytosines are indicated in blue. Guanines are indicated in black. Thymines are indicated in red. Adenines are indicated in green.

b) Alignment between VP1 amino acid sequence (Genbank) and codon-optimized VP1 amino acid sequence from plasmid pMYE100

```
                    10        20        30        40        50        60
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native      1  TTSTGESADPVTATVENYGGETQVQRRQHTDVSFILDRFVKVTPKDQINVLDLMQTPAHT
Optimal     1  TTSTGESADPVTATVENYGGETQVQRRQHTDVSFILDRFVKVTPKDQINVLDLMQTPAHT

                    70        80        90       100       110       120
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native     61  LVGALLRTATYYFADLEVAVKHEGNLTWVPNGAPETALDNTTNPTAYHKAPLTRLALPYT
Optimal    61  LVGALLRTATYYFADLEVAVKHEGNLTWVPNGAPETALDNTTNPTAYHKAPLTRLALPYT

                   130       140       150       160       170       180
               ....|....|....|....|....|....|....|....|....|....|....|....|
Native    121  APHRVLATVYNGNCKYGEGPVTNVRGDLQVLAQKAARTLPTSFNYGAIKATRVTELLYRM
Optimal   121  APHRVLATVYNGNCKYGEGPVTNVRGDLQVLAQKAARTLPTSFNYGAIKATRVTELLYRM

                   190       200       210
               ....|....|....|....|....|....|...
Native    181  KRAETYCPRPLLAIHPSEARHKQKIVAPVKQLL
Optimal   181  KRAETYCPRPLLAIHPSEARHKQKIVAPVKQLL
```

c) Alignment : Positive control VP1 DNA sequence (pMYE100) and VP1 sequence from pGEM®-T-VP1 transformant (for bacterial expression vector pQE-30)

```
                    10        20        30        40        50        60
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        AGGTCCTCACCACCTCCACAGGTGAGTCCGCTGATCCTGTGACTGCCACTGTGGAGAACT
Transformant   GGATC--CACCACCTCCACAGGTGAGTCCGCTGATCCTGTGACTGCCACTGTGGAGAACT

                    70        80        90       100       110       120
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        ACGGTGGTGAGACACAGGTCCAGAGACGCCAGCACACCGATGTCTCCTTCATCTTGGACA
Transformant   ACGGTGGTGAGACACAGGTCCAGAGACGCCAGCACACCGATGTCTCCTTCATCTTGGACA

                   130       140       150       160       170       180
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        GGTTCGTGAAGGTGACACCGAAGGACCAGATCAACGTGTTGGACCTGATGCAGACCCCTG
Transformant   GGTTCGTGAAGGTGACACCGAAGGACCAGATCAACGTGTTGGACCTGATGCAGACCCCTG

                   190       200       210       220       230       240
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        CACACACCTTGGTCGGAGCTCTCCTCCGCACTGCCACCTACTACTTCGCAGATCTCGAGG
Transformant   CACACACCTTGGTCGGAGCTCTCCTCCGCACTGCCACCTACTACTTCGCAGATCTCGAGG

                   250       260       270       280       290       300
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        TGGCAGTGAAGCACGAAGGCAACCTCACCTGGGTCCCTAACGGAGCTCCAGAGACAGCCT
Transformant   TGGCAGTGAAGCACGAAGGCAACCTCACCTGGGTCCCTAACGGAGCTCCAGAGACAGCCT

                   310       320       330       340       350       360
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        TGGACAACACCACCAATCCAACCGCCTACCACAAGGCACCTCTCACCCGGCTCGCACTGC
Transformant   TGGACAACACCACCAATCCAACCGCCTACCACAAGGCACCTCTCACCCGGCTCGCACTGC

                   370       380       390       400       410       420
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        CTTACACCGCACCACACCGCGTCTTGGCCACCGTGTACAACGGCAACTGCAAGTACGGCG
Transformant   CTTACACCGCACCACACCGCGTCTTGGCCACCGTGTACAACGGCAACTGCAAGTACGGCG

                   430       440       450       460       470       480
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        AGGGCCCAGTGACCAACGTGAGAGGTGACCTGCAGGTGCTCGCCCAGAAGGCCGCCAGGA
Transformant   AGGGCCCAGTGACCAACGTGAGAGGTGACCTGCAGGTGCTCGCCCAGAAGGCCGCCAGGA

                   490       500       510       520       530       540
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        CACTGCCTACCTCCTTCAACTACGGAGCCATCAAGGCCACTCGGGTGACTGAGCTCCTCT
Transformant   CACTGCCTACCTCCTTCAACTACGGAGCCATCAAGGCCACTCGGGTGACTGAGCTCCTCT

                   550       560       570       580       590       600
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        ACCGCATGAAGAGGGCCGAGACCTACTGCCCACGGCCTCTCTTGGCCATCCACCCAAGCG
Transformant   ACCGCATGAAGAGGGCCGAGACCTACTGCCCACGGCCTCTCTTGGCCATCCACCCAAGCG

                   610       620       630       640       650
               ....|....|....|....|....|....|..*.|...*|.*..|....|.
Optimal        AGGCCAGACACAAGCAGAAGATCGTGGCTCCTGTGAAGCAGCTCCTCTGAG-GTAC
Transformant   AGGCCAGACACAAGCAGAAGATCGTGGCTCCTGTGAAACAGCTTCTGTAAAAGCTG
```

The optimal (top gene sequence) refers to the codon-optimized VP1 coding sequence within plasmid pMYE100 which served as a positive control and the transformant sequence (bottom gene sequence) refers to the pGEM®-T-VP1 transformant containing restriction enzyme sites 5'-*Bam*HI and 3-'*Hind*III as a stop codon TAA. Restriction enzyme sites are underlined above and the stop codon is shaded in grey. There was a sequencing error with the *Hind*III restriction enzyme site which was confirmed by the functionality of the site in restriction enzyme analysis. Changes in the nucleotide gene sequence are indicated with stars. Cytosines are indicated in blue. Guanines are indicated in black. Thymines are indicated in red. Adenines are indicated in green.

d) Alignment : Positive control VP1 amino acid sequence (pMYE100) to VP1 amino acid sequence from pGEM®-T-VP1 transformant (for bacterial expression vector pQE-30)

```
                    10        20        30        40        50        60
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        TTSTGESADPVTATVENYGGETQVQRRQHTDVSFILDRFVKVTPKDQINVLDLMQTPAHT
Transformant   TTSTGESADPVTATVENYGGETQVQRRQHTDVSFILDRFVKVTPKDQINVLDLMQTPAHT

                    70        80        90       100       110       120
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        LVGALLRTATYYFADLEVAVKHEGNLTWVPNGAPETALDNTTNPTAYHKAPLTRLALPYT
Transformant   LVGALLRTATYYFADLEVAVKHEGNLTWVPNGAPETALDNTTNPTAYHKAPLTRLALPYT

                   130       140       150       160       170       180
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        APHRVLATVYNGNCKYGEGPVTNVRGDLQVLAQKAARTLPTSFNYGAIKATRVTELLYRM
Transformant   APHRVLATVYNGNCKYGEGPVTNVRGDLQVLAQKAARTLPTSFNYGAIKATRVTELLYRM

                   190       200       210
               ....|....|....|....|....|...
Optimal        KRAETYCPRPLLAIHPSEARHKQKIVAPVKQLL
Transformant   KRAETYCPRPLLAIHPSEARHKQKIVAPVKQLL
```

e) Alignment : positive control VP1 gene (pMYE100) and the VP1 gene sequence from pGEM®-T-VP1 transformant (for binary expression vector pMYV497)

```
                    10        20        30        40        50        60
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        GTCCTCACCACCTCCACAGGTGAGTCCGCTGATCCTGTGACTGCCACTGTGGAGAACTAC
Transformant   GGATCCACCACCTCCACAGGTGAGTCCGCTGATCCTGTGACTGCCACTGTGGAGAACTAC

                    70        80        90       100       110       120
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        GGTGGTGAGACACAGGTCCAGAGACGCCAGCACACCGATGTCTCCTTCATCTTGGACAGG
Transformant   GGTGGTGAGACACAGGTCCAGAGACGCCAGCACACCGATGTCTCCTTCATCTTGGACAGG

                   130       140       150       160       170       180
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        TTCGTGAAGGTGACACCGAAGGACCAGATCAACGTGTTGGACCTGATGCAGACCCCTGCA
Transformant   TTCGTGAAGGTGACACCGAAGGACCAGATCAACGTGTTGGACCTGATGCAGACCCCTGCA

                   190       200       210       220       230       240
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        CACACCTTGGTCGGAGCTCTCCTCCGCACTGCCACCTACTACTTCGCAGATCTCGAGGTG
Transformant   CACACCTTGGTCGGAGCTCTCCTCCGCACTGCCACCTACTACTTCGCAGATCTCGAGGTG

                   250       260       270       280       290       300
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        GCAGTGAAGCACGAAGGCAACCTCACCTGGGTCCCTAACGGAGCTCCAGAGACAGCCTTG
Transformant   GCAGTGAAGCACGAAGGCAACCTCACCTGGGTCCCTAACGGAGCTCCAGAGACAGCCTTG

                   310       320       330       340       350       360
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        GACAACACCACCAATCCAACCGCCTACCACAAGGCACCTCTCACCCGGCTCGCACTGCCT
Transformant   GACAACACCACCAATCCAACCGCCTACCACAAGGCACCTCTCACCCGGCTCGCACTGCCT

                   370       380       390       400       410       420
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        TACACCGCACCACACCGCGTCTTGGCCACCGTGTACAACGGCAACTGCAAGTACGGCGAG
Transformant   TACACCGCACCACACCGCGTCTTGGCCACCGTGTACAACGGCAACTGCAAGTACGGCGAG

                   430       440       450       460       470       480
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        GGCCCAGTGACCAACGTGAGAGGTGACCTGCAGGTGCTCGCCCAGAAGGCCGCCAGGACA
Transformant   GGCCCAGTGACCAACGTGAGAGGTGACCTGCAGGTGCTCGCCCAGAAGGCCGCCAGGACA

                   490       500       510       520       530       540
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        CTGCCTACCTCCTTCAACTACGGAGCCATCAAGGCCACTCGGGTGACTGAGCTCCTCTAC
Transformant   CTGCCTACCTCCTTCAACTACGGAGCCATCAAGGCCACTCGGGTGACTGAGCTCCTCTAC

                   550       560       570       580       590       600
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        CGCATGAAGAGGGCCGAGACCTACTGCCCACGGCCTCTCTTGGCCATCCACCCAAGCGAG
Transformant   CGCATGAAGAGGGCCGAGACCTACTGCCCACGGCCTCTCTTGGCCATCCACCCAAGCGAG

                   610       620       630       640       650
               ....|....|....|....|....|....|....|*...|.*..*....|....
Optimal        GCCAGACACAAGCAGAAGATCGTGGCTCCTGTGAAGCAGCTCCTCTGAGGTACC
Transformant   GCCAGACACAAGCAGAAGATCGTGGCTCCTGTGAAACAGCTTCT---GGGTACC
```

The optimal (top gene sequence) refers to the codon-optimized VP1 coding sequence within plasmid pMYE100 which served as a positive control and the transformant sequence (bottom gene sequence) refers to the pGEM®-T-VP1 transformant containing restriction enzyme sites 5'-*Bam*HI and 3'-*Hind*III. Restriction enzyme sites are underlined above. Synonymous changes in the nucleotide gene sequence are indicated with stars. Cytosines are indicated in blue. Guanines are indicated in black. Thymines are indicated in red. Adenines are indicated in green.

f) Alignment: positive control VP1 amino acid sequence (pMYE100) and VP1 amino acid sequence from pGEM®-T-VP1 transformant (for binary expression vector pMYV497)

```
                    10        20        30        40        50        60
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        TTSTGESADPVTATVENYGGETQVQRRQHTDVSFILDRFVKVTPKDQINVLDLMQTPAHT
Transformant   TTSTGESADPVTATVENYGGETQVQRRQHTDVSFILDRFVKVTPKDQINVLDLMQTPAHT

                    70        80        90       100       110       120
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        LVGALLRTATYYFADLEVAVKHEGNLTWVPNGAPETALDNTTNPTAYHKAPLTRLALPYT
Transformant   LVGALLRTATYYFADLEVAVKHEGNLTWVPNGAPETALDNTTNPTAYHKAPLTRLALPYT

                   130       140       150       160       170       180
               ....|....|....|....|....|....|....|....|....|....|....|....|
Optimal        APHRVLATVYNGNCKYGEGPVTNVRGDLQVLAQKAARTLPTSFNYGAIKATRVTELLYRM
Transformant   APHRVLATVYNGNCKYGEGPVTNVRGDLQVLAQKAARTLPTSFNYGAIKATRVTELLYRM

                   190       200       210
               ....|....|....|....|....|...
Optimal        KRAETYCPRPLLAIHPSEARHKQKIVAPVKQLL
Transformant   KRAETYCPRPLLAIHPSEARHKQKIVAPVKQLL
```

REFERENCES

The Huffington Post: Food. 2011. South Korea Reportedly Buries 1.4 Million Pigs Alive To Combat Foot And Mouth Disease. [Online] Available at: http://www.huffingtonpost.com/2011/01/12/south-korea-buries-pigs-alive_n_808119.html?ir=Food

UCSF Chimera--a visualization system for exploratory research and analysis. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. *J Comput Chem.* 2004 Oct;25(13):1605-12.

Abramoff M.D., Magelhaes P.J., Ram S.J. (2004) Image Processing with ImageJ. Biophotonics International 11:36-42.

Abrams C.C., King A.M., Belsham G.J. (1995) Assembly of foot-and-mouth disease virus empty capsids synthesized by a vaccinia virus expression system. J Gen Virol 76:3089-3098.

Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F. (1989) The three-dimensional structure of foot-and-mouth disease virus at 2.9 A resolution. Nature 337:709-716.

Adam Z., Adamska I., Nakabayashi K., Ostersetzer O., Haussuhl K., Manuell A., Zheng B., Vallon O., Rodermel S.R., Shinozaki K., Clarke A.K. (2001) Chloroplast and Mitochondrial Proteases in *Arabidopsis*. A Proposed Nomenclature. Plant Physiol 125:1912-1918.

Adam Z., Clarke A.K. (2002) Cutting edge of chloroplast proteolysis. Trends Plant Sci 7:451-456.

Agterberg M., Adriaanse H., Lankhof H., Meloen R., Tommassen J. (1990) Outer membrane PhoE protein of *Escherichia coli* as a carrier for foreign antigenic determinants: immunogenicity of epitopes of foot-and-mouth disease virus. Vaccine 8:85-91.

Alexandersen S., Mowat N. (2005) Foot-and-mouth disease: host range and pathogenesis. Curr Top Microbiol Immunol 288:9-42.

Alonso M., Hidalgo J., Hendricks L., Velasco A. (1996) Degradation of aggrecan precursors within a specialized subcompartment of the chicken chondrocyte endoplasmic reticulum. Biochem J 316:487-495.

Andrianova E.P., Krementsugskaia S.R., Lugovskaia N.N., Mayorova T.K., Borisov V.V., Eldarov M.A., Ravin N.V., Folimonov A.S., Skryabin1 K.G. (2011) Foot and Mouth Disease Virus Polyepitope Protein Produced in Bacteria and Plants Induces Protective Immunity in Guinea Pigs. Biochemistry (Mosc) 76:339-346.

Arakawa T., Chong D.K.X., Merritt J.L., Langridge W.H.R. (1997) Expression of cholera toxin B subunit oligomers in transgenic potato plants. Transgenic Res 6:403 - 413.

Arntzen C., Plotkin S., Dodet B. (2005) Plant-derived vaccines and antibodies: potential and limitations. Vaccine 23:1753-1756.

Baranowski E., Ruiz-Jarabo C.M., Lim F., Domingo E. (2001) Foot-and-Mouth Disease Virus Lacking the VP1 G-H Loop: The Mutant Spectrum Uncovers Interactions among Antigenic Sites for Fitness Gain. Virology 288:192-202.

Barrett A.J. (1986) The classes of proteolytic enzymes. CRC Press, Boca Raton.

Belsham G.J. (2005) Translation and replication of FMDV RNA. Curr Top Microbiol Immunol 288:43-70.

Benchabane M., Goulet C., Rivard D., Faye L., Gomord V., Michaud D. (2008) Preventing unintended proteolysis in plant protein biofactories. Plant Biotechnol J 6:633-648.

Birch-Machin I., Newell C.A., Hibberd J.M., Gray J.C. (2004) Accumulation of rotavirus VP6 protein in chloroplasts of transplastomic tobacco is limited by protein stability. Plant Biotechnol J 2:261-270.

Broekhuijsen M.P., Blom T., Kottenhagen M., Pouwels P.H., Meloen R.H., Barteling S.J., Enger-Valk B.E. (1986) Synthesis of fusion proteins containing antigenic determinants of foot-and-mouth disease virus. Vaccine 4:119-124.

Brown F. (1992) New approaches to vaccination against foot-and-mouth disease. Vaccine 10:1022-1026.

Buchanan-Wollaston V. (1997) The molecular biology of leaf senescence. J Exp Bot 48:181-199.

Bula C., Wilcox K.W. (1996) Negative effect of sequential serine codons on expression of foreign genes in *Escherichia coli*. Protein Expr Purif 7:92-103.

Callis J. (1995) Regulation of Protein Degradation. Plant Cell 7:845-857.

Canizares M.C., Nicholson L., Lomonossoff G.P. (2005) Use of viral vectors for vaccine production in plants. Immunol Cell Biol 83:263-270.

Carrillo C., Wigdorovitz A., Oliveros J.C., Zamorano P.I., Sadir A.M., Gomez N., Salinas J., Escribano J.M., Borca M.V. (1998) Protective immune response to foot-and-mouth disease virus with VP1 expressed in transgenic plants. J Virol 72:1688-1690.

Carrillo C., Wigdorovitz A., Trono K., Dus Santos M.J., Castanon S., Sadir A.M., Ordas R., Escribano J.M., Borca M.V. (2001) Induction of a virus-specific antibody response to foot and mouth disease virus using the structural protein VP1 expressed in transgenic potato plants. Viral Immunol 14:49-57.

Cazzonelli C.I., Velten J. (2006) An in vivo, luciferase-based, *Agrobacterium*-infiltration assay system: implications for post-transcriptional gene silencing. Planta 224:582-597.

Chen G.-F.T., Inouye M. (1990) Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. Nucleic Acids Res 18:1465-1473.

Choi J.H., Lee S.Y. (2004) Secretory and extracellular production of recombinant proteins using *Escherichia coli*. Appl Microbiol Biotechnol 64:625-635.

Christensen A.H., Quail P.H. (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res 5:213-218.

Christensen A.H., Sharrock R.A., Quail P.H. (1992) Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. Plant Mol Biol 18:675-689.

Conrad U., Fiedler U. (1998) Compartment-specific accumulation of recombinant immunoglobulins in plant cells: an essential tool for antibody production and immunomodulation of physiological functions and pathogen activity. Plant Mol Biol 38:101–109.

Corchero J.L., Viaplana E., Benito A., Villaverde A. (1996) The position of the heterologous domain can influence the solubility and proteolysis of [beta]-galactosidase fusion proteins in *E. coli*. J Biotechnol 48:191-200.

D'Aoust M.-A., Lavoie P.-O., Belles-Isles J., Bechtold N., Martel M., Vézina L.-P. (2009) Transient expression of antibodies in plants using syringe agroinfiltration. Recombinant Proteins From Plants, pp. 41-50.

Daniell H., Chebolu S., Kumar S., Singleton M., Falconer R. (2005) Chloroplast-derived vaccine antigens and other therapeutic proteins. Vaccine 23:1779-1783.

Daniell H., Khan M.S., Allison L. (2002) Milestones in chloroplast genetic engineering: an environmentally friendly era in biotechnology. Trends Plant Sci 7:84-91.

Delannoy M., Alves G., Vertommen D., Ma J., Boutry M., Navarre C. (2008) Identification of peptidases in *Nicotiana tabacum* leaf intercellular fluid. Proteomics 8:2285-2298.

Demirevska K., Simova-Stoilova L., Fedina I., Georgieva K., Kunert K. (2010) Response of Oryzacystatin I Transformed Tobacco Plants to Drought, Heat and Light Stress. Journal of Agronomy & Crop Science 196:90-99.

Doran P.M. (2006) Foreign protein degradation and instability in plants and plant tissue cultures. Trends Biotechnol 24:426-432.

Drenth J., Jansonius J.N., Koekoek R., Marrink J., Munnik J., Wolthers B.G. (1962) The crystal structure of papain C: I. Two-dimensional fourier syntheses. J Mol Biol 5:398-407.

Dus Santos M.J., Carrillo C., Ardila F., Rios R.D., Franzone P., Piccone M.E., Wigdorovitz A., Borca M.V. (2005) Development of transgenic alfalfa plants containing the foot and mouth disease virus structural polyprotein gene P1 and its utilization as an experimental immunogen. Vaccine 23:1838-1843.

Dus Santos M.J., Wigdorovitz A. (2005) Transgenic plants for the production of veterinary vaccines. Immunol Cell Biol 83:229-238.

Dus Santos M.J., Wigdorovitz A., Trono K., Rios R.D., Franzone P.M., Gil F., Moreno J., Carrillo C., Escribano J.M., Borca M.V. (2002) A novel methodology to develop a foot and mouth disease virus (FMDV) peptide-based vaccine in transgenic plants. Vaccine 20:1141-1147.

Dwek R.A. (1996) Glycobiology: Toward Understanding the Function of Sugars. Chem Rev 96:683-720.

Ebel J., Mithofer A. (1998) Early events in the elicitation of plant defence. Planta 206:335–348.

Ellgaard L., Helenius A. (2003) Quality control in the endoplasmic reticulum. Nat Rev Mol Cell Biol 4:181.

Everett K.M., Dickey L., Parsons J., Loranger R., Wingate V. (2012) Development of a plant-made pharmaceutical production platform. Bioprocess Int 10:16-26.

Fan S.-G., Wu G.-J. (2005) Characteristics of plant protease inhibitors and their applications in combating phytophagous insects. Bot Bull Acad Sinica 46:273-292.

Faye L., Boulaflous A., Benchabane M., Gomord V., Michaud D. (2005) Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. Vaccine 23:1770-1778.

Feliu J.X., Villaverde A. (1998) Engineering of solvent-exposed loops in *Escherichia coli* β-galactosidase. FEBS Letters 434:23-27.

Filho J.X. (1992) The biological roles of serine and cysteine proteinase inhibitors in plants. R Bras Fisiol Veg 4:1-6.

Fox G., Parry N.R., Barnett P.V., McGinn B., Rowlands D.J., Brown F. (1989) The Cell Attachment Site on Foot-and-Mouth Disease Virus Includes the Amino Acid Sequence RGD (Arginine-Glycine-Aspartic Acid). J Gen Virol 70:625-637.

Fritschy J.-M. (2008) Is my antibody-staining specific? How to deal with pitfalls of immunohistochemistry. Eur J Neurosci 28:2365-2370.

Fry E.E., Stuart D.I., Rowlands D.J. (2005) The structure of foot-and-mouth disease virus. Curr Top Microbiol Immunol 288:71-101.

Gatz C., Frohberg C., Wendenburg R. (1992) Stringent repression and homogeneous de-repression by tetracycline of a modified CaMV 35S promoter in intact transgenic tobacco plants. Plant J 2:397-404.

Gennaro M.L., Greenaway P.J. (1983) Nucleotide sequences within the cholera toxin operon. Nucleic Acids Res 11:3855-3861.

Georgiou G., Segatori L. (2005) Preparative expression of secreted proteins in bacteria: status report and future prospects. Curr Opin Biotechnol 16:538-545.

Gils M., Kandzia R., Marillonnet S., Klimyuk V., Gleba Y. (2005) High-yield production of authentic human growth hormone using a plant virus-based expression system. Plant Biotechnol J 3:613-620.

Gomord V., Faye L. (2004) Posttranslational modification of therapeutic proteins in plants. Curr Opin Plant Biol 7:171-181.

Gomord V., Fitchette A.-C., Menu-Bouaouiche L., Saint-Jore-Dupas C., Plasson C., Michaud D., Faye L. (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. Plant Biotechnol J 8:564 - 587.

Gomord V., Sourrouille C., Fitchette A.-C., Bardor M., Pagny S., Lerouge P., Faye L. (2004) Production and glycosylation of plant-made pharmaceuticals: the antibodies as a challenge. Plant Biotechnol J 2:83-100.

Goulet C., Benchabane M., Anguenot R., Brunelle F., Khalf M., Michaud D. (2010) A companion protease inhibitor for the protection of cytosol-targeted recombinant proteins in plants. Plant Biotechnol J 8:142–154.

Goulet C., Khalf M., Sainsbury F., D'Aoust M.-A., Michaud D. (2011) A protease activity–depleted environment for heterologous proteins migrating towards the leaf cell apoplast. Plant Biotechnol J:1-12.

Griffin J.F.T. (2002) A strategic approach to vaccine development: animal models, monitoring vaccine efficacy, formulation and delivery. Adv Drug Deliv Rev 54:851-861.

Gruber V., Berna P.P., Arnaud T., Bournat P., Clément C., Mison D., Olagnier B., Philippe L., Theisen M., Baudino S., Bénicourt C., Cudrey C., Bloës C., Duchateau N., Dufour S., Gueguen C., Jacquet S., Ollivo C., Poncetta C., Zorn N., Ludevid D., Van Dorsselaer A., Verger R., Doherty A., Mérot B., Danzin C. (2001) Large-scale production of a therapeutic protein in transgenic tobacco plants: effect of subcellular targeting on quality of a recombinant dog gastric lipase. Mol Breed 7:329-340.

Grubman M.J., Baxt B. (2004) Foot-and-Mouth Disease. Clin Microbiol Rev 17:465–493.

Hammond-Kosack K.E., Jones J.D.G. (1996) Resistance gene dependent plant defence responses. Plant Cell 8:1773-1791.

He D.M., Qian K.X., Shen G.F., Li Y.N., Zhang Z.F., Su Z.L., Shao H.B. (2007) Stable expression of foot-and-mouth disease virus protein VP1 fused with cholera toxin B subunit in the potato (*Solanum tuberosum*). Colloids Surf B Biointerfaces 55:159-163.

Hellebust H., Murby M., Abrahmsen L., Uhlen M., Enfors S.-O. (1989) Different Approaches to Stabilize a Recombinant Fusion Protein. Nat Biotech 7:165-168.

Hellens R., Mullineaux P., Klee H. (2000) Technical Focus : A guide to *Agrobacterium* binary Ti vectors. Trends Plant Sci 5:446-451.

Hellwig S., Drossard J., Twyman R.M., Fischer R. (2004) Plant cell cultures for the production of recombinant proteins. Nat Biotech 22:1415-1422.

Henry D. (2006) Production of biopharmaceuticals and vaccines in plants via the chloroplast genome. Biotechnol J 1:1071-1079.

Herman E., Schmidt M. (2004) Endoplasmic Reticulum to Vacuole Trafficking of Endoplasmic Reticulum Bodies Provides an Alternate Pathway for Protein Transfer to the Vacuole. Plant Physiol 136:3440-3446.

Hoorn R.A.L.V.d., Laurent F., Roth R., Wit P.J.G.M.D. (2000) Agroinfiltration is a versatile tool that facilitates comparative analyses of Avr9/Cf-9-Induced and Avr4/Cf-4-induced necrosis. Phytopathology 13:439 - 446.

Ichikawa Y., Yamagata H., Tochikubo K., Udaka S. (1993) Very efficient extracellular production of cholera toxin B subunit using *Bacillus brevis*. FEMS Microbiol Lett 111:219-224.

Inoue H., Nojima H., Okayama H. (1990) High efficiency transformation of *Escherichia coli* with plasmids. Gene 96:23-28.

James E., Lee J. (2001) The production of foreign proteins from genetically modified plant cells plant cells. In: Zhong, J., Byun, S., Cho, G., Choi, J., Haigh, J., Honda, H., James, E., Kijne, J., Kim, D., Kobayashi, T., Lee, J., Kino-oka, M., Linden, J., Liu, C., Memelink, J., Mirjalili, N., Nagatome, H., Taya, M., Phisaphalong, M., van der Heijden, R., Verpoorte, R. (Eds.). Springer Berlin / Heidelberg, pp. 127-156.

Janssen B.-J., Gardner R. (1990) Localized transient expression of GUS in leaf discs following cocultivation with *Agrobacterium*. Plant Mol Biol 14:61-72.

Jiang L., Li Q., Li M., Zhou Z., Wu L., Fan J., Zhang Q., Zhu H., Xu Z. (2006) A modified TMV-based vector facilitates the expression of longer foreign epitopes in tobacco. Vaccine 24:109-115.

Johansen L.K., Carrington J.C. (2001) Silencing on the Spot. Induction and Suppression of RNA Silencing in the *Agrobacterium*-Mediated Transient Expression System. Plant Physiol 126:930-938.

Jones N.H. (1977) Fracastoro and Henle: A re-appraisal of their contribution to the concept of communicable diseases. Med Hist 21:61-68.

Kang T.J., Han S.C., Jang M.O., Kang K.H., Jang Y.S., Yang M.S. (2004) Enhanced Expression of B-Subunit of *Escherichia coli* Heat-Labile Enterotoxin in Tobacco by Optimization of Coding Sequence. Appl Biochem Biotechnol 117:175 - 187.

Kapila J., De Rycke R., Van Montagu M., Angenon G. (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. Plant Sci J 122:101-108.

Kay R., Chan A., Daly M., McPherson J. (1987) Duplication of CaMV 35S Promoter Sequences Creates a Strong Enhancer for Plant Genes. Science 236:1299-1302.

Knowles N.J., Samuel A.R. (2003) Molecular epidemiology of foot-and-mouth disease virus. Virus Res 91:65-80.

Kupper H., Keller W., Kurz C., Forss S., Schaller H., Franze R., Strohmaier K., Marquardt O., Zaslavsky V.G., Hofschneider P.H. (1981) Cloning of cDNA of major antigen of foot and mouth disease virus and expression in E. coli. Nature 289:555-559.

Kurz C., Forss S., Kupper H., Strohmaier K., Schaller H. (1981) Nucleotide sequence and corresponding amino acid sequence of the gene for the major antigen of foot and mouth disease virus. Nucleic Acids Res 9:1919-1931.

Laemmli U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680 - 685.

Leippert M., Beck E., Weiland F., Pfaff E. (1997) Point mutations within the betaG-betaH loop of foot-and-mouth disease virus O1K affect virus attachment to target cells. J Virol 71:1046-1051.

Lentz E., Segretin M., Morgenfeld M., Wirth S., Dus Santos M., Mozgovoj M., Wigdorovitz A., Bravo-Almonacid F. (2010) High expression level of a foot and mouth disease virus epitope in tobacco transplastomic plants. Planta 231:387-395.

Li Y., Sun M., Liu J., Yang Z., Zhang Z., Shen G. (2006) High expression of foot-and-mouth disease virus structural protein VP1 in tobacco chloroplasts. Plant Cell Rep 25:329-333.

Light A., Frater R., Kimmel J.R., Smith E.L. (1964) Current Status of the Structure of Papain: The Linear Sequence, Active Sulfhydryl Group, and the Disulfide Bridges. Proceedings of the National Academy of Sciences of the United States of America 52:1276-1283.

Loeffler F., Frosch P. (1897) Summarischer Bericht ueber der Ergebnisse der Untersuchungen zur Erforschung der Maul-und-Klauenseuche. Zent. Bakt Parasitkde Abt. I 22:257–259.

Loeffler F., Frosch P. (1898) Report of the Commission for Research on foot-and-mouth disease. Zent. Bakt. Parasitkde. Abt.I 23:371–391.

Ma J.K., Drake P.M., Christou P. (2003) The production of recombinant pharmaceutical proteins in plants. Nat. Rev. Genet. 4:794–805.

Ma J.K., Hein M.B. (1995) Plant antibodies for immunotherapy. Plant Physiol 109:341-346.

Ma J.K.C., Drake P.M.W., Chargelegue D., Obregon P., Prada A. (2005) Antibody processing and engineering in plants, and new strategies for vaccine production. Vaccine 23:1814-1818.

Mahy B.W. (2005a) Foot-and-mouth disease virus. Curr Top Microbiol Immunol 288:1-178.

Mahy B.W. (2005b) Introduction and history of foot-and-mouth disease virus. Curr Top Microbiol Immunol 288:1-8.

Makrides S. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. Microbiol. Rev. 60:512-538.

Maliga P. (2002) Engineering the plastid genome of higher plants. Curr Opin Plant Biol 5:164-172.

Maniatis T., Fritsch E.F., Sambrook J. (1982) Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Mason H.S., Lam D.M., Arntzen C.J. (1992) Expression of hepatitis B surface antigen in transgenic plants. Proceedings of the National Academy of Sciences of the United States of America 89:11745-11749.

McConn M., Creelman R.A., Bell E., Mullet J.E., Browse J. (1997) Jasmonate is essential for insect defense in Arabidopsis. Proc. Natl Acad. Sci. USA 94:5473–5477.

McKenna T.S., Lubroth J., Rieder E., Baxt B., Mason P.W. (1995) Receptor binding site-deleted foot-and-mouth disease (FMD) virus protects cattle from FMD. J. Virol. 69:5787-5790.

Menkhaus T.J., Bai Y., Zhang C., Nikolov Z.L., Glatz C.E. (2004) Considerations for the Recovery of Recombinant Proteins from Plants. Biotechnol Prog 20:1001-1014.

Mikschofsky H., König P., Keil G.M., Hammer M., Schirrmeier H., Broer I. (2009) Cholera toxin B (CTB) is functional as an adjuvant for cytoplasmatic proteins if directed to the endoplasmatic reticulum (ER), but not to the cytoplasm of plants. Plant Sci 177:35-42.

Morita A., Umemura T.-a., Kuroyanagi M., Futsuhara Y., Perata P., Yamaguchi J. (1998) Functional dissection of a sugar-repressed α-amylase gene (RAmy1A) promoter in rice embryos. FEBS Lett 423:81-85.

Munro S., Pelham H.R.B. (1987) A C-terminal signal prevents secretion of luminal ER proteins. Cell 48:899-907.

Nakamura K., Matsuoka K. (1993) Protein Targeting to the Vacuole in Plant Cells. Plant Physiol. 101:1-5.

Napier J.A., Richard G., Shewry P.R. (1998) Trafficking and Stability of Heterologous Proteins in Transgenic Plants. Recombinant Proteins from Plants, pp. 189-202.

Neubauer P., Hofmann K., Holst O., Mattiasson B., Kruschke P. (1992) Maximizing the expression of a recombinant gene in *Escherichia coli* by manipulation of induction time using lactose as inducer. Appl Microbiol Biotechnol 36:739-744.

Neuhaus J.-M., Rogers J.C. (1998) Sorting of proteins to vacuoles in plant cells. Plant Mol Biol 38:127-144.

Nuttall J., Vine N., Hadlington J.L., Drake P., Frigerio L., Ma J.K.C. (2002) ER-resident chaperone interactions with recombinant antibodies in transgenic plants. Eur J Biochem 269:6042-6051.

Odell J.T., Nagy F., Chua N.-H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. Nature 313:810-812.

Oem J.K., Lee K.N., Cho I.S., Kye S.J., Park J.H., Joo Y.S. (2004) Comparison and analysis of the complete nucleotide sequence of foot-and-mouth disease viruses from animals in Korea and other PanAsia strains. Virus Genes 29:63-71.

Oliveira A.S., Filho J.X., Sales M.P. (2003) Cysteine proteinases and cystatins. Brazilian archives of Biology and Technology 46:91 - 104.

Pagny S., Lerouge P., Faye L., Gomord V. (1999) Signals and mechanisms for protein retention in the endoplasmic reticulum. J Exp Bot 50:157–164.

Pan L., Zhang Y.G., Wang Y.L., Wang B.Q., Xie Q.G. (2006) Protective immune response of guinea pigs against challenge with foot and mouth disease virus by immunization with foliar extracts from transgenic tomato plants expressing the FMDV structural protein VP1. Wei Sheng Wu Xue Bao 46:796-801.

Pérez-Filgueira D.M., Brayfield B.P., Phiri S., Borca M.V., Wood C., Morris T.J. (2004) Preserved antigenicity of HIV-1 p24 produced and purified in high yields from plants inoculated with a tobacco mosaic virus (TMV)-derived vector. J Virol Methods 121:201-208.

Perlak F.J., Deaton R.W., Armstrong T.A., Fuchs R.L., Sims S.R., Greenplate J.T., Fischhoff D.A. (1990) Insect Resistant Cotton Plants. Nat Biotech 8:939-943.

Petruccelli S., Otegui M.S., Lareu F., Dinh O.T., Fitchette A.-C., Circosta A., Rumbo M., Bardor M., Carcamo R., Gomord V., Beachy R.N. (2006) A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. Plant Biotechnol J 4:511-527.

Prins A., van Heerden P.D., Olmos E., Kunert K.J., Foyer C.H. (2008) Cysteine proteinases regulate chloroplast protein content and composition in tobacco leaves: a model for dynamic interactions with ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) vesicular bodies. J Exp Bot 59:1935-1950.

Pueyo J.J., Chrispeels M.J., Herman E.M. (1995) Degradation of transport-competent destabilized phaseolin with a signal for retention in the endoplasmic reticulum occurs in the vacuole. Planta 196:586-596.

Ratish G. V.S., Suryanarayana V.V. (1999) C-terminal region of VP1 of selected foot-and-mouth disease virus serotypes: expression in *E. coli* and affinity purification Acta Virol 43:205-211.

Rice J., Ainley W.M., Shewen P. (2005) Plant-made vaccines: biotechnology and immunology in animal health. Anim Health Res Rev 6:199-209.

Rivard D., Anguenot R., Brunelle F., Le V.Q., Vézina L.-P., Trépanier S., Michaud D. (2006) An in-built proteinase inhibitor system for the protection of recombinant proteins recovered from transgenic plants. Plant Biotechnol J 4:359-368.

Robinson D.G., Oliviusson P., Hinz G. (2005) Protein Sorting to the Storage Vacuoles of Plants: A Critical Appraisal. Traffic 6:615-625.

Rodríguez M., Ramírez N.I., Ayala M., Freyre F., Pérez L., Triguero A., Mateo C., Selman-Housein G., Gavilondo J.V., Pujol M. (2005) Transient expression in tobacco leaves of an aglycosylated recombinant antibody against the epidermal growth factor receptor. Biotechnol Bioeng 89:188-194.

Rogowsky P.M., Close T.J., Chimera J.A., Shaw J.J., Kado C.I. (1987) Regulation of the vir genes of *Agrobacterium tumefaciens* plasmid pTiC58. J Bacteriol 169:5101-5112.

Ruhlman T., Ahangari R., Devine A., Samsam M., Daniell H. (2007) Expression of cholera toxin B - proinsulin fusion protein in lettuce and tobacco chloroplasts - oral administration protects against development of insulitis in non-obese diabetic mice. Plant Biotechnol J 5:495-510.

Sahdev S., Khattar S., Saini K. (2008) Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. Mol Cell Biochem 307:249-264.

Sainsbury F., Liu L., Lomonossoff G.P. (2009) Cowpea mosaic virus-based systems for the expression of antigens and antibodies in plants. Methods Mol Biol 483:25-39.

Sainsbury F., Lomonossoff G. (2008) Extremely high-level and rapid transient protein production in plants without the use of viral replication. Plant Physiol 148:1212 - 1218.

Samuel A.R., Knowles N.J. (2001) Foot-and-mouth disease type O viruses exhibit genetically and geographically distinct evolutionary lineages (topotypes). J Gen Virol 82:609-621.

Schaller A. (2004) A cut above the rest: the regulatory function of plant proteases. Planta 220:183–197.

Schatz G., Dobberstein B. (1996) Common Principles of Protein Translocation Across Membranes. Science 271:1519-1526.

Schöb H., Kunz C., Meins Jr F. (1997) Silencing of transgenes introduced into leaves by agroinfiltration: a simple, rapid method for investigating sequence requirements for gene silencing. Mol Gen Genet 256:581-585.

Shire S.J., Bock L., Ogez J., Builder S., Kleid D., Moore D.M. (1984) Purification and immunogenicity of fusion VP1 protein of foot and mouth disease virus. Biochemistry 23:6474-6480.

Smalle J., Vierstra R.D. (2004) The ubiquitin 26S proteasome proteolytic pathway. Annu Rev Plant Biol 55:555-590.

Solomon M., Belenghi B., Delledonne M., Menachem E., Levine A. (1999) The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. Plant Cell 11:431-444.

Song H., Fang W., Wang Z., Zheng D., Du J., Li H., Li Y., Qiu B. (2004a) Detection of foot-and-mouth virus antibodies using a purified protein from the high-level expression of codon-optimized, foot-and-mouth disease virus complex epitopes in *Escherichia coli*. Biotechnol Lett 26:1277-1281.

Song H., Zhou L., Fang W., Li Y., Wang X., Fang H., Li X., Wu M., Qiu B. (2004b) High-level expression of codon optimized foot-and-mouth disease virus complex epitopes and cholera toxin B subunit chimera in *Hansenula polymorpha*. Biochem Biophys Res Commun 315:235-239.

Sriraman R., Bardor M., Sack M., Vaquero C., Faye L., Fischer R., Finnern R., Lerouge P. (2004) Recombinant anti-hCG antibodies retained in the endoplasmic reticulum of transformed plants lack core-xylose and core-α(1,3)-fucose residues. Plant Biotechnol J 2:279-287.

Stevens L.H., Stoopen G.M., Elbers I.J.W., Molthoff J.W., Bakker H.A.C., Lommen A., Bosch D., Jordi W. (2000) Effect of climate conditions and plant developmental stage on the stability of antibodies expressed in transgenic tobacco. Plant Physiol. 124:173-182.

Stoger E., Ma J.K.C., Fischer R., Christou P. (2005) Sowing the seeds of success: pharmaceutical proteins from plants. Curr Opin Biotechnol 16:167-173.

Streatfield S.J. (2007) Approaches to achieve high-level heterologous protein production in plants. Plant Biotechnol J 5:2-15.

Streatfield S.J., Lane J.R., Brooks C.A., Barker D.K., Poage M.L., Mayor J.M., Lamphear B.J., Drees C.F., Jilka J.M., Hood E.E., Howard J.A. (2003) Corn as a production system for human and animal vaccines. Vaccine 21:812-815.

Strohmaier K., Franze R., Adam K.H. (1982) Location and Characterization of the Antigenic Portion of the FMDV Immunizing Protein. J Gen Virol 59:295-306.

Su C., Duan X., Wang X., Wang C., Cao R., Zhou B., Chen P. (2007) Heterologous expression of FMDV immunodominant epitopes and HSP70 in *P. pastoris* and the subsequent immune response in mice. Vet Microbiol 124:256-263.

Suryanarayana V.V.S., Viswanathan S., Ratish G., Bist P., Prabhudas K., Gajendragad M.R., Natarajan C. (1999) *E. coli* expressed proteins as diagnostic reagents for typing of foot-and-mouth disease virus. Archives of Virology 144:1701-1712.

Svergun D.I., Petoukhov M.V., Koch M.H.J. (2001) Determination of Domain Structure of Proteins from X-Ray Solution Scattering. Biophys J 80:2946-2953.

Terpe K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl Microbiol Biotechnol 72:211-222.

Torre J.L.L., Grubman M.J., Baxt B., Bachrach H.L. (1980) The structural polypeptides of aphthovirus are phosphoproteins. Microbiology 77:7444 - 7447.

Toyofuku K., Umemura T.-a., Yamaguchi J. (1998) Promoter elements required for sugar-repression of the RAmy3D gene for α-amylase in rice. FEBS Lett 428:275-280.

Triguero A., Cabrera G., Rodríguez M., Soto J., Zamora Y., Pérez M., Wormald M.R., Cremata J.A. (2011) Differential N-glycosylation of a monoclonal antibody expressed in tobacco leaves with and without endoplasmic reticulum retention signal apparently induces similar in vivo stability in mice. Plant Biotechnol J 9:1120-1130.

Turk V., Bode W. (1991) The cystatins: Protein inhibitors of cysteine proteinases. FEBS Lett 285:213-219.

Twyman R.M., Stoger E., Schillberg S., Christou P., Fischer R. (2003) Molecular farming in plants: host systems and expression technology. Trends Biotechnol 21:570-578.

Ueda T., Seo S., Ohashi Y., Hashimoto J. (2000) Circadian and senescence-enhanced expression of a tobacco cysteine protease gene. Plant Mol Biol 44:649-657.

Usha R., Rohll J.B., Spall V.E., Shanks M., Maule A.J., Johnson J.E., Lomonossoff G.P. (1993) Expression of an animal virus antigenic site on the surface of a plant virus particle. Virology 197:366-374.

van der Hoorn R.A.L. (2008) Plant Proteases: From Phenotypes to Molecular Mechanisms. Annu Rev Plant Biol 59:191-223.

Van der Vyver C., Schneidereit J., Driscoll S., Turner J., Kunert K., Foyer C.H. (2003) Oryzacystatin I expression in transformed tobacco produces a conditional growth phenotype and enhances chilling tolerance. Plant Biotechnol J 1:101-112.

van Lierop M.J., van Maanen K., Meloen R.H., Rutten V.P., de Jong M.A., Hensen E.J. (1992) Proliferative lymphocyte responses to foot-and-mouth disease virus and three FMDV peptides after vaccination or immunization with these peptides in cattle. Immunology 75:406-413.

Vaquero C., Sack M., Chandler J., Drossard J., Schuster F., Monecke M., Schillberg S., Fischer R. (1999) Transient expression of a tumor-specific single-chain fragment and a chimeric antibody in tobacco leaves. Proceedings of the National Academy of Sciences 96:11128-11133.

Verdaguer N., Mateu M.G., Andreu D., Giralt E., Domingo E., Fita I. (1995) Structure of the major antigenic loop of foot-and-mouth disease virus complexed with a neutralizing antibody: direct involvement of the Arg-Gly-Asp motif in the interaction. Embo J 14:1690-1696.

Vierstra R.D. (1996) Proteolysis in plants: mechanisms and functions. Plant Mol Biol 32:275-302.

Vierstra R.D. (2003) The ubiquitin/26S proteasome pathway, the complex last chapter in the life of many plant proteins. Trends Plant Sci 8:135-142.

Vitale A., Denecke J. (1999) The Endoplasmic Reticulum - Gateway of the Secretory Pathway. The Plant Cell Online 11:615-628.

Voinnet O., Rivas S., Mestre P., Baulcombe D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J 33:949–956.

von Heijne G. (1990) The signal peptide. J Membr Biol 115:195-201.

Wandelt C.I., Khan M.R.I., Craig S., Schroeder H.E., Spencer D., Higgins T.J.V. (1992) Vicilin with carboxy-terminal KDEL is retained in the endoplasmic reticulum and accumulates to high levels in the leaves of transgenic plants. Plant J 2:181-192.

Wang K. (2006) Agrobacterium protocols. Totowa, N.J. , Humana.

Westers L., Westers H., Quax W.J. (2004) *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. Biochim Biophys Acta (BBA) - Molecular Cell Research 1694:299-310.

Wigdorovitz A., Carrillo C., Dus Santos M.J., Trono K., Peralta A., Gómez M.C., Ríos R.D., Franzone P.M., Sadir A.M., Escribano J.M., Borca M.V. (1999a) Induction of a protective antibody response to foot and mouth disease virus in mice following oral or parenteral immunization with alfalfa transgenic plants expressing the viral structural protein VP1. Virology 255:347-353.

Wigdorovitz A., Perez Filgueira D.M., Robertson N., Carrillo C., Sadir A.M., Morris T.J., Borca M.V. (1999b) Protection of mice against challenge with foot and mouth disease virus (FMDV) by immunization with foliar extracts from plants infected with recombinant tobacco mosaic virus expressing the FMDV structural protein VP1. Virology 264:85-91.

Wroblewski T., Tomczak A., Michelmore R. (2005) Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis*. Plant Biotechnol J 3:259-273.

Wydro M., Kozubek E., Lehmann P. (2006) Optimization of transient *Agrobacterium*-mediated gene expression system in leaves of *Nicotiana benthamiana*. Acta Biochim Pol 53:289-300.

Yang Y., Li R., Qi M. (2000) *In vivo* analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. Plant J 22:543-551.

Yin J., Li G., Ren X., Herrler G. (2007) Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. J Biotechnol 127:335-347.

Zambryski P., Joos H., Genetello C., Leemans J., Montagu M.V., Schell J. (1983) Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. The EMBO J 2:2143-2150.

Zamorano P., Wigdorovitz A., Chaher M.T., Fernandez F.M., Carrillo C., Marcovecchio F.E., Sadir A.M., Borca M.V. (1994) Recognition of B and T cell epitopes by cattle immunized with a synthetic peptide containing the major immunogenic site of VP1 FMDV 01 Campos. Virology 201:383-387.